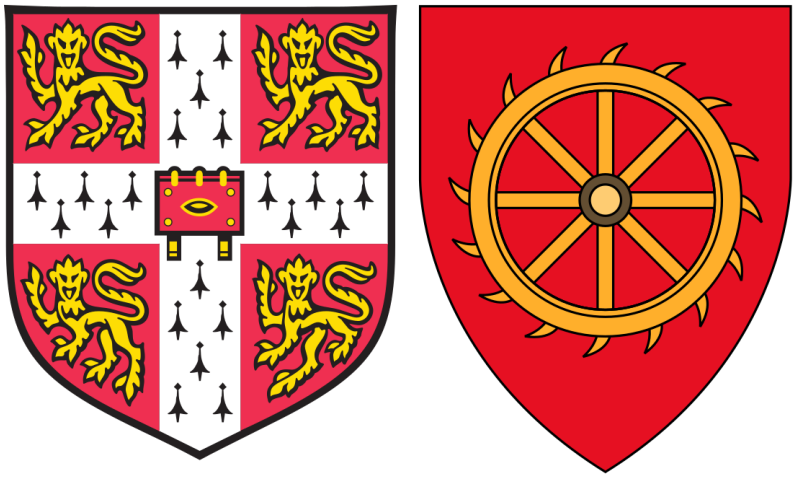
# Phage host range and definition of genes implicated in Type III toxin-antitoxin-mediated abortive infection

**Ray Chai**



**St Catharine's College**

**University of Cambridge**

**Department of Biochemistry**

**This dissertation is submitted for the degree of Doctor of Philosophy**

**September 2018**

**Phage host range and definition of genes implicated in Type III toxin-antitoxin-mediated abortive infection**

Ray Chai

## Summary

Bacteria are under constant threat by their viral parasites, the bacteriophages (phages) and have evolved a range of anti-phage systems to defend themselves. One of these systems is termed abortive infection (Abi) where, upon phage infection, an Abi system may be activated which initiates a bacteriostatic or bactericidal response. While the infected bacteria do not obviously benefit from the activation of these systems, the cessation of bacterial growth or premature cellular suicide prevents the release of phage progeny. Thus Abi can be viewed as an altruistic process as only the remaining clonal bacterial population benefits. The Type III toxin-antitoxin systems have previously been shown to be involved in Abi, however the mechanisms through which these systems are activated are still poorly understood. A common approach to reveal the phage product involved in triggering these systems is to first determine the mutations that a previously sensitive phage evolves to escape after exposure to an Abi system. Analysis of viral "escape" mutants has been used in this study to try to elucidate the activation mechanism(s) of two Type III systems ($ToxIN_{Pa}$ and $TenpIN_{Pl}$) of several environmental phages. Several new phage products were identified in escape mutants as candidate factors involved in circumventing Abi – and possible roles in phage metabolism predicted. Furthermore, the genomes of several phages that could not evolve escapes, or were insensitive to Abi, are sequenced and these data exposed interesting curiosities regarding Abi (as well as the discovery of several novel and rare phages). Previously, no coliphage was identified that was capable of escape of the $ToxIN_{Pa}$ or $TenpIN_{Pl}$ systems. However, this study defined and characterised the first $ToxIN_{Pa}$ and $TenpIN_{Pl}$ coliphage escapes as well as a new method for isolating host-dependent coliphage escapes. Finally, multiple phages that infect the insect pathogen *Photorhabdus luminescens* TT01 (the bacterial strain from which the $TenpIN_{Pl}$ system originated) were isolated, genomically sequenced and characterised in terms of host range. The results revealed a large superfamily of flagellum-dependent phages that exhibit remarkable host promiscuity, possibly defining the most promiscuous phages thus far identified.

## Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the biology degree committee.

Ray Chai
September 2018

Dans les champs de l’observation le hasard ne favorise que les esprits prepares

Louis Pasteur, December 1854

**Acknowledgements**

I would like to thank my supervisor Professor George Salmond for his support and guidance during my time in his lab. His enthusiasm for phage biology was infectious and I sincerely wished I had more time in the lab. I have always felt free to discuss and carry out any experiment I wanted and many of the serendipitous findings during this study were a result of this.

I am very grateful to Rita Monson who has helped me out countless times within and outside the lab and for also constantly reminding me that “everything is going to ok”. I am thankful also for Alison Rawlinson who taught me how to make my first phage lysate and for keeping the lab running smoothly. Thank you also to Chin Mei Lee who has always given sound scientific advice and for being my “lab mum”.

Special thanks must go to my BBM, Alice Ahn. “Banter bay” was never quite the same after you left. I have always remembered my first year very fondly and have missed our constant debates. In this regard I must also thank Catharine Fan and I will always remember the ridiculous treasure hunt for Mr Gold orchestrated by you. Mr Gold has since watched over me till the very last day and I hope he brings good luck to his next owner as he has for me.

Thanks to Bihe Chen, Andy Day, Alex Quintero, Audrey Crousilles and Eve Maunders. It has been a fun journey to start our PhDs together and I wish you all the best to where life takes you next. Thanks to postdocs Nathalie Goeders and Jessica Bergman for both giving excellent scientific advice. Thanks also to all past and present Salmond and Welch lab members that I have had the pleasure to work alongside with.

There were many collaborators that I have had the pleasure of working with during this PhD and sadly I regret that I could not include all the data that resulted from these works. Thanks to Del Pickard for gifting us the YSD1 phage and carrying out experiments with the TT01 phages. Thanks to Gabriel De Freitas Almeida for taking beautiful helium-ion microscopy images for the TT01 phage experiments which sadly could not be included in this dissertation. Thank you also to David Clark for gifting us all the bacterial insect pathogens for the TT01 host range tests. I am also very thankful to Jason Clark and the rest of the Fixed phage team. I very much enjoyed carrying out

the “phage fishing” experiments and am very regretful that I did not have more time to perform more experiments and to include it in this dissertation. I am also very grateful for the opportunity to teach phage biology to Parkside school. Chapter 5 was the serendipitous result of this collaboration.

I would also like to thank past Salmond lab members Pete Fineran and Tim Blower, true legends of the Salmond lab phage team. Pete has also been a pleasure to talk to about phage and I have appreciated all the new insights after our chats. Tim has also been very supportive and I have to thank him for teaching me how to make “proper” scientific figures.

This PhD would not have happened if it weren’t for the support of my parents. They have given me this privilege, for which I am eternally grateful. Finally, I would like to thank Ji, my best friend who I married in the final year of my PhD. We’ve come a long way from meeting as undergraduate freshers all those years ago in Bristol and it still amazes me how much we have grown since then. Your commitment and successes of your own PhD has been a constant inspiration and I am extremely proud to be your husband.

# Table of Contents

## List of Figures

## List of Tables

**Abbreviations**

| | |
|---|---|
| Abi | abortive infection |
| ACT | Artemis Comparison Tool |
| Ap | ampicillin |
| BLAST | basic local alignment search tool |
| Bp | base pair(s) |
| Bt | *Bacillus thuringiensis* |
| Cm | Chloramphenicol |
| CptI | CptN inhibitor |
| CptN | *Coprococcus* Type III toxin |
| CRISPR | clustered regularly interspaced short palindromic repeat |
| $dH_20$ | deionised water |
| (g)DNA | (genomic) deoxyribonucleic acid |
| EOP | efficiency of plating |
| EtOH | ethanol |
| Er | *Eubacterium rectale* |
| FS | frameshift |
| HEPES | N-(2-hydroxylethyl)-piperazine-N'-(2-ethanesulfonic acid) |
| HF | High fidelity |
| LB (agar) | lysogeny broth (agar) |
| Km | kanamycin |
| MB(A) | marine broth (agar) |
| MES | 2-(N-morpholino) ethanesulfonic acid |
| MOI | multiplicity of infection |
| OD600 | optical density at 600 nm |
| ORF | open reading frame |
| Pa | *Pectobacterium atrosepticum* |
| PAGE | polyacrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| Pfu | plaque forming unit(s) |
| Phage(s) | bacteriophage(s) |
| Pl | *Photorhabdus luminescens* |
| RNA | ribonucleic acid |
| SCRI | Scottish Crop Research Institute |

| | |
|---|---|
| SDS | sodium dodecyl sulphate |
| TA | toxin-antitoxin |
| TAE | tris-acetate-EDTA |
| TEM | transmission electron microscopy |
| TenpI | TenpN inhibitor |
| TenpN | toxin endogenous to *Photorhabdus* toxin |
| ToxI | ToxN inhibitor |
| ToxN | toxin |
| w/v | weight per volume |
| Wt | wild type |

# Chapter 1 Introduction

## 1.1 Bacteriophages

### 1.1.1 History and significance

Bacteriophages (phages) “bacteria eaters” are the viral parasites of bacteria (D’Herelle, 2007). The discovery of phages is somewhat controversial and credited to two different researchers who may have discovered them independently (Duckworth, 1976). The first description of phages has been attributed to Frederick Twort, a British physician who carried out experiments while he was superintendent at the Brown Institution at London University. While working with the vaccinia virus he used an unfiltered fluid that was used in small pox vaccinations to try and replicate the virus. The vaccinia virus could not replicate but he noted that micrococci grew on the agar plate instead. He noticed that a number of the micrococci colonies showed a physical manifestation of a “disease” and described these colonies as glassy and transparent. Furthermore, he noted that by touching a non-diseased *i.e.* non glassy colony with a little of a glassy colony, the glassy phenotype spread throughout the colony as it grew. The agent that lead to the glassy phenotype still retained activity after filtration and so could be replicated. These observations suggested that he was in fact working with a bacterial virus, however, Twort never explicitly stated that. He cautiously suggested that his observations had not disproven the agent to be a virus and could be caused by other alternatives such as a self-replicating enzyme. Unfortunately, he was not able to carry this research further as he states at the end of the paper, “I regret that financial considerations have prevented my carrying these researches to a definite conclusion, but I have indicated the lines along which others more fortunately situated can proceed” (Twort, 1915).

The “more fortunately situated” turned out to be Felix d’Herelle, a French-Canadian microbiologist who, while working at the Pasteur institute in France, characterised “clear spots” he would occasionally find on lawns of bacteria after mixing filtered faecal samples from patients with dysentery with cultures of dysentery bacilli. In these experiments he suggested that the causation of these “clear spots” was from an invisible microbe as he could not see the microbe through light microscopy (D’Herelle, 2007). He called these microbes “bacteriophages” *i.e.* bacteria eater where the word phage comes from the Greek work “phagein” meaning to eat. It became apparent to

him that this bacterial parasite might be used to treat bacterial infections in humans paving the way for phage therapy (Fruciano and Bourne, 2007; Summers, 2001). It was initially heralded as an Erlich's "magic bullet" due to its specificity in killing bacterial hosts. However, these early phage trials are often criticised for the lack of scientific rigor and mixed results (Fruciano and Bourne, 2007). Furthermore, with the discovery of antibiotics and the association of phage therapy with the Soviet Union at the time (phage therapy was often referred to as Stalin's cure) the interest in phage research for therapy declined, at least in the West (Stone, 2002; Wittebole et al., 2014).

At this point in time, interest in bacteriophage research was diminishing however this was all about to change. In 1937, Max Delbrück a German physicist turned biologist, moved to Caltech to work on *Drosophila* with T. H Morgan (Keen, 2015). As a trained physicist he was looking to find the simplest model organism with which to study the basic principles of biology. *Drosophila* proved to be too complex, however during his time at Caltech, he met with Emory Ellis a postdoctoral researcher, who was working with phages at the time. Delbruck soon became fascinated by phages and choose to use them as his model organism instead of flies. He began to collaborate with other scientists including, Salvador Luria and Alfred Hershey and others. This marked the birth of the Phage Group, which was a group of researchers using phages as a model "organism" to study the fundamental principles of biology. The research performed by members of the Phage Group defined some of the most important in the field of biology (**Table 1**).

Phages have had a tremendous impact on our understanding of the fundamental principles of biology (Keen, 2015). A list of several of the most significant experiments involving phages is shown in **Table 1**. Furthermore, studies of phages have also led to the development of many molecular tools such as phage display, phage typing, restriction enzymes and more recently the CRISPR-Cas systems. Advances in sequencing has led to a boom in new phage sequences published which has revealed just how diverse these entities are and just how little of the phage world is understood. The majority of published phage genomes encode many proteins of unknown function potentially providing a rich source of novel biology.

**Table 1.** Key phage experiments from the 1940-70

* signifies a Nobel laureate. Members of the phage group are in bold. This table was taken from (Keen, 2015).

| Year | Researcher(s) | Finding(s) |
|---|---|---|
| 1943 | **S. Luria***, **M. Delbrück*** | Mutation is spontaneous and random (pre-adaptive), not directed (post-adaptive) |
| 1945 | **S. Luria*** | Viruses can mutate so as to be able to successfully infect previously resistant bacteria |
| 1946 | **A. Hershey***, **M. Delbrück*** | Two viruses can genetically recombine when co-infecting the same cell |
| 1950 | A. Lwoff*, L. Siminovitch, N. Kjeldgaard | Ultraviolet radiation can induce the excision of a viral genome from that of its host |
| 1951 | V. Freeman | Viruses can contribute to bacterial virulence (*e.g.,* by encoding diphtheria toxin) |
| 1952 | **A. Hershey***, M. Chase | DNA, not protein, is unambiguously the hereditary material of life |
| 1952 | N. Zinder, J. Lederberg* | Viruses can transfer DNA between cells (transduction) |
| 1952 | **S. Luria***, M. Human | Once adapted to a particular host, viruses' ability to infect other hosts is greatly diminished (this "restriction" was later shown to result from restriction enzymes) |
| 1955 | **S. Benzer** | The sequence of a gene is linear; recombination can occur between adjacent nucleotides |
| 1961 | B. Hall, S. Spiegelman | Complementary DNA and RNA can hybridize |
| 1961 | F. Crick*, L. Barnett, **S. Brenner***, R. Watts-Tobin | Nucleotides are read three units at a time (as codons) to form proteins |
| 1961 | **S. Brenner***, F. Jacob*, **M. Meselson** | mRNA is the intermediate between DNA and protein; mRNA is translated into protein by ribosomes |
| 1962 | D. Nathans*, G. Notani, J. Schwartz, N. Zinder | Purified RNA can direct the synthesis of its encoded protein in a cell-free environment |
| 1967 | M. Gellert (and other laboratories) | DNA ligase joins together DNA fragments |
| 1967 | M. Goulian, A. Kornberg*, R. Sinsheimer | DNA can be synthesized from its precursors *in vitro* |
| 1969 | R. Okazaki, T. Okazaki, K. Sakabe, K. Sugimoto, A. Sugino | Lagging strand DNA synthesis is discontinuous |

### 1.1.2 Bacteriophage taxonomy

Phage taxonomy has not been as straight-forward partly due to the fact that there is no single protein present in all phages and so phage taxonomy is constantly changing. Originally, d'Herelle suggested that all phages were a single species with multiple races which he called *Bacteriophagum intestinale* (D'Herelle, 1918). Then with the invention of the electron microscope, phages could finally be observed and were classified based on their morphology (Luria et al., 1943). A greater understanding of the nature of nucleic acids also showed that the genetic material of phages differed and could be spilt into either single stranded DNA, double stranded DNA, single stranded RNA and double stranded RNA (Ackermann, 2009). In 1966, the International committee for the taxonomy of viruses (ICTV) was formed which aimed to classify all viruses, including phages. The ICTV created a taxonomy tree based on the nature of the nucleic acid of phages and their morphology. For example, the order *Caudovirales* contains all tailed phages which was originally spilt into three different families, *Myoviridae*, *Siphoviridae* and *Podoviridae*, based on the tail morphology. Phages belonging to *Myoviridae*, *Siphoviridae*, *Podoviridae* have long contractile tails, long non contractile tails, and short non contractile tails, respectively. These phage families are then broken down further into subfamilies and genera based on other factors such as the hosts they were isolated from, sizes of genomes and whether the genome was linear, circular or supercoiled (Nelson, 2004).

There are obvious short-comings to the methods chosen for classfication by the ICTV as it does not take into account the genetic or proteomic information of each phage. This gives the false impression that phages within a family are more evolutionary related to each other than those outside. This is not always the case, for example, the podovirus P22 has a short non contractile tail and does not share homology with many other *Podoviridae* but shares homology with the *Siphoviridae* member λ. Both P22 and λ phages share such high genetic homology that recombination experiments between the genomes of both phages were able to produce some functional hybrids (Botstein and Herskowitz, 1974). However, P22 and λ are still classified as being in different families instead of being grouped together. Many alternative approaches to improving phage taxonomy have been proposed based on using genetic and/or proteomic data (Nelson, 2004). The ICTV is trying to acknowledge these problems and is constantly updating taxonomy. Recently, *Caudovirales*, the tailed phages that make up 96% of all sequenced phages have been spilt into four families. The new family called

*Ackermannviridae* was originally part of the *Myoviridae* family as members contained long contractible tails but, as they showed little genetic homology to other *Myoviridae*, the new family was proposed (Kuhn et al., 2017). It is likely that in coming years the original three *Caudovirales* familes, *Myoviridae*, *Podoviridae* and *Siphoviridae* will be subdivided even further (Aiewsakun et al., 2018).

### 1.1.3 Bacteriophage lifecycles

Phages are obligate intracellular parasites and so require a bacterial host in order to replicate. They have been shown to have at least four different replication life cycles (**Figure 1.1**). These are the lytic, lysogenic, pseudolysogenic and chronic cycles (Weinbauer, 2004).

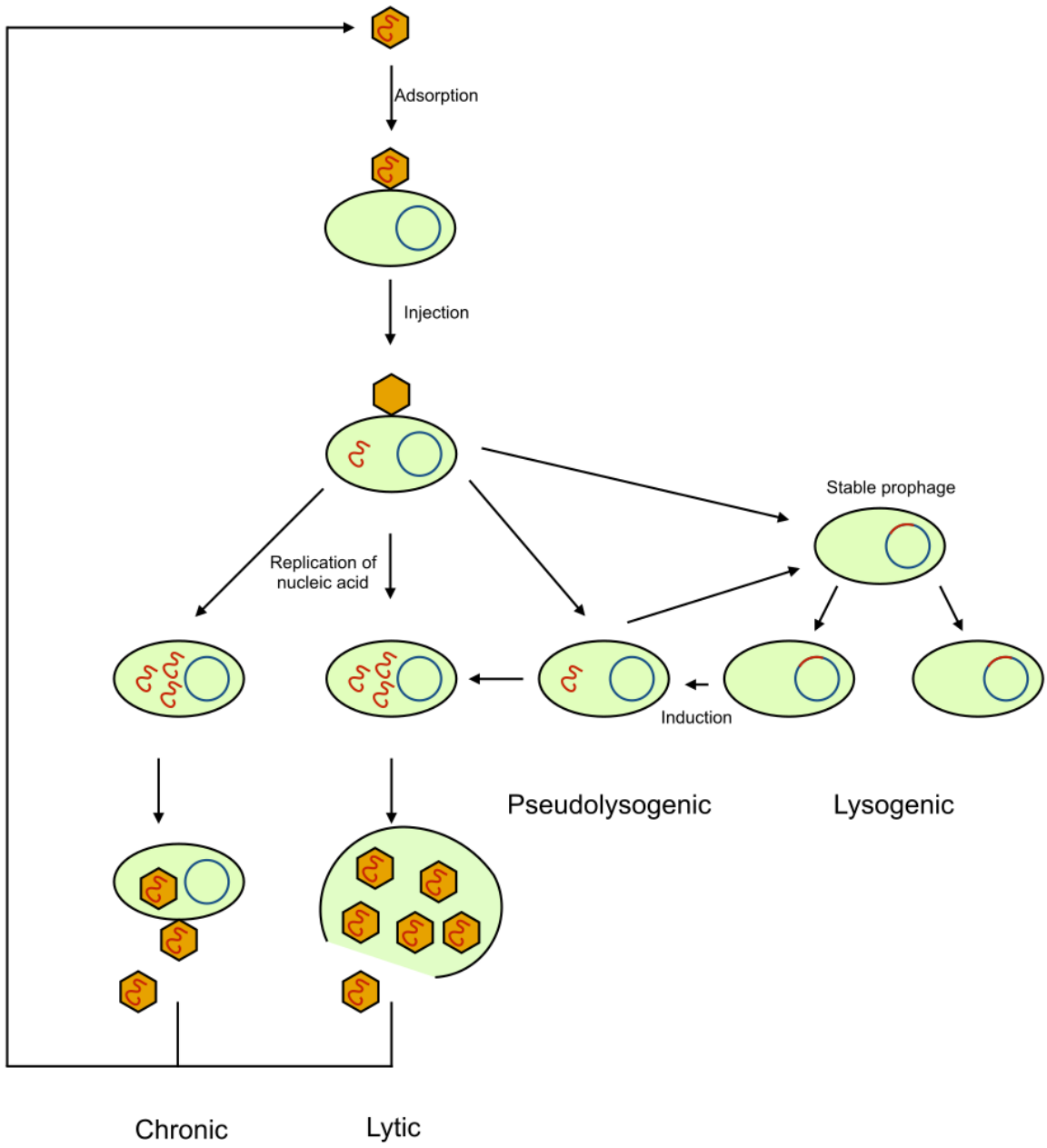



**Figure 1.1** Phage replication life cycles

Simplified phage replication life cycles showing the four different replication cycles. Figure adapted from (Weinbauer, 2004).

### 1.1.3.1 Lytic cycle

The lytic cycle leads to the death of the host cell that the phage invades and is most commonly observed in the *Caudoviridae* order. The first step of a lytic cycle is called adsorption where the phage makes contact with a specific receptor located on the surface of the bacterial cell. Common bacterial receptors include lipopolysaccharide (LPS), flagella and a range of outer membrane proteins (Bertozzi Silva et al., 2016). The recognition of these receptors by the phage can be through tail fibres or tail spike proteins. After adsorption the next step involves injection, when the genomic material is injected through the tail of the phage into the bacterial host. Phage genetic material is then replicated using the host transcription machinery. For the *Tevenvirinae*, including T4, phage genes are under the control of three different promoters depending on when they need to be transcribed. The genes are labelled as early, middle and late and are sequentially transcribed (Miller et al., 2003). The early genes code for proteins involved in the metabolism of the phage and includes RNA and DNA polymerases, nucleases etc. During the early phase, the phage genomic material is replicated and the conditions are set to favour middle gene transcription. The middle genes code for proteins involved in the structural aspect of the phage such as the major capsid protein, tail tube protein and tail fibre. The phage particles are assembled at this stage and finally late gene products are made. Late genes encode phage products involved in the lysis of the host, such as holins and endolysins. Production of these phage products leads to the lysis of the host cell, releasing many newly formed phage progeny into the surrounding environment to start the cycle anew.

### 1.1.3.2 Lysogeny

The lytic lifecycle causes the death of the host which means that the phage progeny now have to find a new host in which to replicate. This can be disadvantageous if potential hosts are scarce and the environment is not favourable. Lysogeny is a life cycle that has evolved in some phages and does not always lead to the lysis of the host cell. Phages that undergo lysogeny are called temperate phages and upon injecting their genomic material into their host can incorporate it via recombination into the chromosomes of their hosts where they reside as prophages. For a few phages, such as P1, the prophage exists as a plasmid (Łobocka et al., 2004). While their hosts replicate, the genomic material of the temperate phages is also replicated. In this way all further bacterial progeny will carry phages without causing lysis of the host. Bacteria

carrying these temperate phages are known as lysogens. Temperate phages no longer need to constantly locate a new host in which to replicate and are protected from the environment. Lysogens can also benefit by carrying prophages as they are immune to further later infection of the same type of phage, via superinfection systems (discussed in detail later). The prophages themselves carry many new genes that can be beneficial to lysogens in a symbiotic relationship. However, prophages can also enter a lytic cycle. This can happen spontaneously in a small number of cells, for example in phage λ lysogens this occurs at $10^{-8}$ to $10^{-5}$ per cell (Czyz et al., 2001). There are various factors that can induce the release of prophages. One of the common causes is DNA damage, in fact, phage λ was discovered accidently by Esther Lederberg by exposing a λ lysogen to UV light (Gottesman and Weisberg, 2004; Lederberg and Lederberg, 1953). Other effectors include antibiotics, hydrogen peroxide, mitomycin C, changes in nutrients and uptake of foreign DNA (Banks et al., 2003; Choi et al., 2010; Czyz et al., 2001; Schuster et al., 1973; Setlow et al., 1973). Prophages can also cause the induction of other prophages carried by the host. An example of this is the exploitation of phage P2 by phage P4 in a process described as molecular piracy. P2 encodes an anti-repressor called Cox which causes the induction of P4 allowing it effectively “steal” the structural proteins encoded by P2, forming empty P2 virions and packaging its own DNA in these virions (Christie and Dokland, 2012). On the other hand *Enterococcus* prophages namely, pp1, pp3 and pp5 have also been shown to prevent the induction of prophages pp4 and pp6 (Matos et al., 2013).

Lysogeny can offer advantages for both phages and their hosts. For instance, temperate phages often carry useful genes including those involved in antibiotic resistance and virulence (Fortier and Sekulovic, 2013; Larrañaga et al., 2018). Many temperate phages are also transducers *i.e.* they can accidently package host genetic material into their virions and after lysis of the host, inject it into a new host, thus acting as a means for horizontal gene transfer (Schicklmaier et al., 1998). As lysogens are immune to further infection by the same phage other susceptible bacteria can be lysed by the spontaneous release of phage, thus reducing the competition as well as providing nutrients from the lysed cells (Harrison and Brockhurst, 2017). Prophages are also known to code for bacteriocins which can be produced by the lysogen to kill competitors (Gebhart et al., 2012). While carrying prophages can seem like a double edge sword for many lysogens, prophages can also be domesticated by losing genes

necessary for the formation of viral particles making them unable to enter the lytic cycle (Strathern and Herskowitz, 1975).

#### 1.1.3.3 Pseudolysogeny

Pseudolysogeny is an intermediate between lytic and lysogeny cycles and is a phenomenon that is still not well understood. There are many definitions of pseudolysogeny because it may be a broad term used to describe a range of different life cycles exhibited by phages. Generally speaking, pseudolysogeny is when a phage infects a bacterial cell but stops replicating, likely due to host conditions being unfavourable. The genetic material of the phage does not stably recombine with the host genome as in many temperate phages and does not replicate. Instead it can remain unchanged until favourable growth conditions allow the resident phage DNA to either undergo a lytic or lysogenic cycle (Clokie et al., 2011). Nutrient starvation has been linked to pseudolysogeny hence, once host cells are no longer starved, phages may be able to then be capable of continuing either a lytic or lysogenic lifecycle (Ripp and Miller, 1998).

#### 1.1.3.4 Chronic

The release of phage particles from infected bacterial cells does not always cause the lysis of the host cell. Filamentous phages can undergo a chronic lifecycle whereby, instead of lysing the bacterial cell, the viral particles are constantly shed by the cell without causing death to the bacterial host (Mai-Prochnow et al., 2015). This lifecycle has been exploited technically since bacterial hosts are not killed and filamentous phages (such as M13 or fd) can be engineered for phage display applications (Sidhu, 2001).

## 1.2 Anti-phage mechanisms

Bacteria might be the most abundant organisms on Earth, however, phages represent the most abundant biological entities outnumbering their bacterial hosts 10:1 (Brüssow and Hendrix, 2002). There is an estimated $10^{31}$ phages in the world (Brüssow and Hendrix, 2002). This number is likely an underestimate as the methodologies used to determine the number of phages have various limitations. The original group that predicted that there was $2.5x10^{8}$ viral particles per ml of natural water (which is likely where the $10^{31}$ figure was extrapolated from) counted plaque forming units using various host bacteria (Bergh et al., 1989). The limitations of these experiments are

obvious as a vast majority of phages would remain unaccounted for. Other techniques that have been employed are epifluorescence microscopy using either DAPI or SYBR green, both dyes mainly used for dsDNA quantification (Noble and Fuhrman, 1998; Suttle et al., 1990). Pulsed field gel electrophoresis has also been used to measure viral populations in the sea, however, this again only takes account of dsDNA viruses (Steward et al., 2000). To address this issue flow cytometry has been used which is able to count RNA viruses as well but flow cytometry counts still underestimate the number of viral particles present (Tomaru and Nagasaki, 2007). Many ssDNA and RNA viruses are still thought to be missed by these techniques and may be more abundant than previously thought (Breitbart, 2011; Steward et al., 2013). Nonetheless, the number of predicted phage infections is still predicted to be $10^{25}$ per second leading to a huge global turnover in bacterial populations (Lima-Mendez et al., 2007). Phages and bacteria have existed for millions of years so it is hardly surprising that, to survive constant phage onslaught, bacteria have developed multiple anti-phage mechanisms. At every stage of the phage infection cycle, bacteria have evolved mechanisms to control infection (**Figure 1.2**). These mechanisms are described in further detail in the following sections.

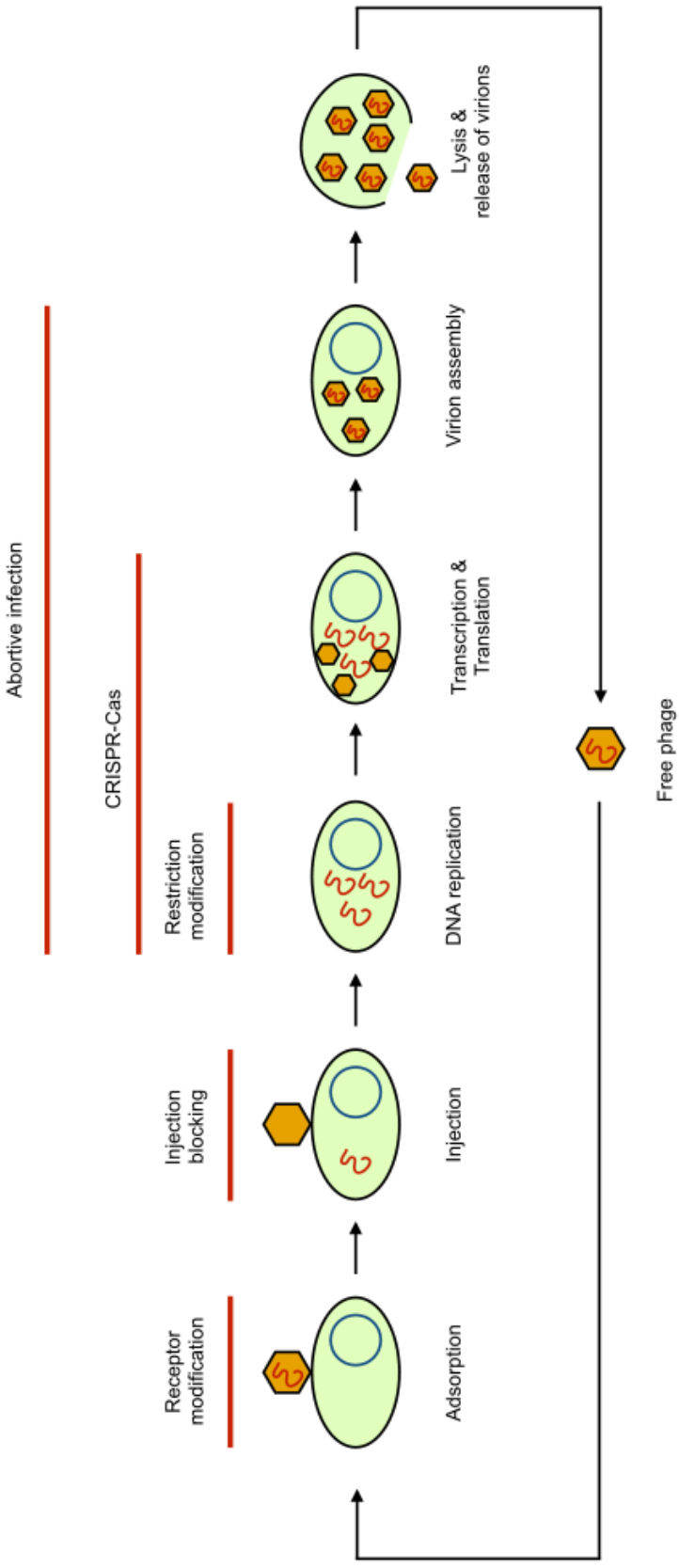



**Figure 1.2** Bacteria have evolved anti-phage systems at every stage of the phage infection cycle The red bars show the stage of the life cycles at which the various anti-phage systems occur.

### 1.2.1 Receptor modification

The first step of phage infection is called adsorption whereby the phage recognises a specific receptor on the bacterial surface, and phages tend to be very selective with their receptors. Consequently, should the bacteria alter this receptor in some way it may prevent the phage from recognising it and so prevent an infection. One problem that bacteria have is that many of the common phage receptors are important components of the bacteria such as LPS, flagella or outer membrane proteins. Alteration of such receptors may lead to phage resistance however, it can lead to a fitness cost for the bacteria which may be unfavourable in the long term, particularly if the phage population is relatively low. However, specific mutations can occur in some outer membrane protein genes that prevent phage adsorption but still retain outer membrane protein biological function (Morona et al., 1984). Another method to prevent phages recognising the receptor is for bacteria to reduce expression of genes that upregulate the receptor or alternatively to completely prevent access to the receptor by masking it in extracellular matrix (Hammad, 1998; Kim et al., 2015; Tan et al., 2015). Phages have been shown to circumvent this by either evolving to recognise an alternate receptor or by destroying the extracellular matrix masking the receptor by integrating depolymerising enzymes in their capsids (Kim et al., 2015). *E. coli* mutants have been shown to down regulate the LamB protein, the receptor for λ. Correspondingly, phage λ can evolve to circumvent this by recognising an alternate receptor, the OmpF protein (Meyer et al., 2012).

### 1.2.2 Superinfection

Once adsorbed onto a suitable receptor the phage has to inject its genomic material into its host. The injection step can be prevented by a mechanism called a superinfection system. These systems originated from temperate phages to prevent competition from other phages within an already infected host. As these systems are encoded on many prophages, some bacterial strains can already be immune to a range of phages. Many superinfection systems are known but the mechanisms involved are still not well understood, although they are all membrane associated. Examples of superinfection systems include *Streptococcus thermophilus* phage TP-J34 where the superinfection control system is a membrane associated lipoprotein that prevents the injection of *Siphoviridae* phages by interaction with their tape measure

proteins (Bebeacua et al., 2013). The tape measure protein of *Siphoviridae* phages acts as a channel in which DNA can bypass the membrane and so is required for injection. The most well-known superinfection control systems are two in the phage T4. The Imm system prevents the injection of DNA by modification of the injection site while the Sp system inactivates phage lysozyme (Lu and Henning, 1989, 1994).

### 1.2.3 Restriction and modification systems

If a phage manages to inject its nucleic acid into the bacterial host it can be targeted by restriction and modification (RM) systems. These are by far the most studied of anti-phage systems due to their exploitation in molecular biology (Roberts, 2005). The phenomena of RM systems were first observed in 1952 by Luria and Human and in 1953 by Bertani and Weigle (Bertani and Weigle, 1953; Luria and Human, 1952). It was noted that the efficiency with which phages infected certain hosts was dependent on the host on which they were previously propagated. It appeared that phages remembered what host they had been propagating on so that while they replicated efficiency in the previous host, infection of a new host would be less efficient. The phages would then replicate efficiently in this new host but poorly in the previous host however, efficiency to the original host could be restored by propagating the phage back on it. Werner Arber proved that this phenomenon was the result of restriction modification systems which consisted of two different enzymes (Arber, 1965). The restriction enzyme acted as a pair of molecular scissors cleaving specific DNA sequences. The modification enzyme prevents the restriction enzyme from cleaving the bacterial DNA by modifying it by methylation. RM systems have been characterised in four types based on their mechanism and genetic organisation (Loenen et al., 2014).

Phages have evolved a range of strategies to cope with RM systems. Occasionally the phage DNA is also methylated by the methylase enzyme leading to all future phage progeny DNA being methylated and insensitive to the RM system. However, should a phage infect a host with a different methylase enzyme the progeny DNA follows this new modification and as the phage DNA no longer carries the same modification as the original host it reverts back to a sensitive state (Labrie et al., 2010). Phages can also modify their own DNA to prevent restriction. Phage T4 contains hydroxymethylcytosine in place of cytosine which makes it insensitive to all RM

systems that recognise any sequences with cytosine (Miller et al., 2003). As phages modify their DNA, bacteria have also developed systems to specifically recognise these modifications. Modification-dependent systems specifically recognise modified DNA, examples include the *Streptococcus pneumoniae* DpnI system as well as the McrA, McrBC and Mrr systems of *E. coli* (Raleigh and Wilson, 1986; Vovis and Lacks, 1977). Phage T4 is resistant to modification dependent systems as its hydromethlated cytosines are also glucosylated. However, strain CT596 of *E. coli* contains a prophage which carries the GmrS-GmrD system that can recognise and cut glucosylated DNA (Bair and Black, 2007). Furthermore, T4 has responded in the form of a small nuclease inhibitor, IPI. During the injection stage of T4, hundreds of IPI proteins are injected alongside the genome which are able to directly inhibit the GmrS-GmrD system (Rifat et al., 2008). This “molecular arms race” does not end here as certain strains of *E. coli* have a GmrS-GmrD system that is encoded on a single protein and are immune to IPI. This is evidence of the Red Queen hypothesis which theorises that every organism has to constantly evolve and adapt lest they are outcompeted by other organisms that are also perpetually evolving and thus face extinction. It shows how both phages and their bacterial hosts are continuously evolving complex solutions to combat each other (Stern and Sorek, 2011).

### 1.2.4 CRISPR-Cas

The CRISPR-Cas systems are the most specific antiphage defence identified thus far and are often described as an adaptive immune system in bacteria (Barrangou et al., 2007). CRISPR-Cas is short for clustered regularly interspaced short palindromic repeats and CRISPR associated proteins. The CRISPR array is a genetic locus made out of short repeating sequences that are separated by non-identical DNA sequences called spacers (Jansen et al., 2002). These spacers contain DNA originating from plasmids or phage genomes (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). Spacers are acquired during a phage infection, at the adaptation stage of the CRISPR-Cas system. Once the spacer has been acquired the next stage, expression, begins with the transcription of the entire CRISPR locus into a single RNA, pre-crRNA. This pre-crRNA is then cut and processed by Cas protein into mature crRNA. Finally, interference occurs during the next phage infection when identical phage sequence is recognised by the mature crRNA and targeted for degradation (Rath et al., 2015). There are three established CRISPR-Cas types defined by the composition of the Cas

proteins and the differences in the mechanism of acquisition and priming of spacers (S. Makarova et al., 2011). There are also CRISPR-Cas systems that do not fit in any of the three types which will likely mean the number of different types of CRISPR-Cas systems will increase as more is understood about them (S. Makarova et al., 2011). Due their ability to be exploited as precise genome editing tools, CRISPR-Cas systems have become a heavily researched topic, again highlighting how bacteriophage research has led to revolutionary molecular biological tools (Sander and Joung, 2014).

The CRISPR Cas systems have been found in the genomes of 40% of bacteria and 90% of archaea (Sorek et al., 2008). Although these figures may be lower if considering uncultivable microorganisms (Burstein et al., 2016). Nonetheless, there are examples of phages evolving counter measures to these systems. Phage T4 is likely already partially resistant to CRISPR-Cas as its glucosylated hydroxymethlated genome may interfere with the CRISPR-Cas system cutting its genome (Tao et al., 2018). Perhaps the simplest solution to escaping the CRISPR-Cas system are through point mutations of the phage genome in the spacer sequence as it would prevent recognition of the phage genome, and have the least impact on the phage (Deveau et al., 2008). More complex solutions have also been observed as phages have evolved a variety of anti-CRISPR proteins (Hynes et al., 2018; Zhu et al., 2018). Phages have also turned the CRISPR-Cas system against their bacteria host. A group of *Vibrio cholerae* phages encode their own CRISPR-Cas system that targets another anti-phage system in their *V. cholerae* hosts (Seed et al., 2013).

### 1.2.5 Abortive infection

Abortive infection, the focus of this thesis is used to describe a wide range of anti-phage systems acting after the injection of phage nucleic acid. The first reported description of abortive infection was in 1952 when phage T2r a mutant of phage T2 (selected as it produces larger plaques) showed a reduced replication efficiency i.e lower multiplies of infection on starved *E. coli* strain B (Benzer, 1952). The scientist who performed the study, Seymour Benzer, called this phenomenon abortive adsorption. As a side note, Benzer, a member of the “phage group” eventually worked on the *rIIA* and *rIIB* genes and, using crosses of *rII* mutants produced the first fine genomic structure map (Benzer, 1959; Jayaraman, 2008). His most famous rII mutant, r1589, was used by Crick et al to show that DNA was in codons consisting of three

nucleotides (Crick et al., 1961). The work on abortive adsorption was characterised further by Samson Gross who once again observed that phage T2rI, another T2 phage variant, replicated less efficiently in starved *E. coli* K12 (Gross, 1954). Gross was unsure of the phage replication stage when this inhibition occurred and renamed it as abortive infection.

One of the key hallmarks of an abortive infection is that the phage infected cell does not directly benefit from it. When an abortive infection system is activated by a sensitive phage it can have a bacteriostatic or bactericidal effect on the host. While halting growth, or initiating the premature death, of the cell does not benefit the bacterium, this leads to a reduced number, or no, phage progeny being released into the surrounding environment. The result is that the surrounding clonal bacteria are protected from further infection and so abortive infection is often viewed as a last resort “altruistic” response (Dy et al., 2014a; Shub, 1994).

Abortive infection (Abi) is a post-infection control system and it encompasses a wide range of different anti-phage systems, all with very different mechanisms. Abi has been observed in a wide range of bacteria including *E. coli*, *Lactococcus lactis*, *Vibrio cholerae*, *Shigella*, *Bacillus* and *Pectobacterium atrosepticum* (Chopin et al., 2005; Chowdhury et al., 1989; Fineran et al., 2009; Smith et al.; Snyder and McWilliams, 1989; Tran et al., 1999). One of the first Abi systems identified was the RexAB system found in phage λ (Snyder and McWilliams, 1989). Wild type *E. coli* K12 is a lysogen carrying phage λ as a prophage, which allows it to be resistant to other coliphages. Phage T4, however, is unaffected by the RexAB system if it contains a functional *r*II locus i.e. a functional RIIA and RIIB protein. Thus T4 *r*II mutants are unable to plaque on *E. coli* K12 (Landsmann et al., 1982). This was first seen by Benzer and investigated further by him and others (Benzer, 1959; Crick et al., 1961). The RexAB system consists of two proteins; RexA and RexB. Activation of the system in phage T4 is triggered by RexA acting as a sensor which detects a phage protein-bacterial DNA complex (Parma et al., 1992). Once detected, two RexA proteins cause the activation of RexB, a membrane protein that acts as an ion channel causing membrane depolarisation and, as membrane potential is lost, the death of the cell (Parma et al., 1992). However, whether this leads to an altruistic suicide protecting the bacterial population remains controversial as there is also evidence suggesting that protection conferred by the RexAB system may be through creation of an unsuitable environment

for phage replication (Slavcev and Hayes, 2002). Thus the enhanced survival of infected cells, rather than death, may protect the bacterial population.

While T4 is immune to wild type K12 lysogen abortive infection, overexpression of *rexAB* can still abort T4 suggesting that levels of *rexAB* are important for activity (Shinedling et al., 1987). However, mutations in the *motA* gene expression are able to make T4 insensitive again; the mechanism of which is not understood (Shinedling et al., 1987). Furthermore, overexpression of *rexA* halts cellular growth and causes a decrease in intracellular ATP levels similar to abortive infection phages (Snyder and McWilliams, 1989). This may suggest that perturbation of the stoichiometry of RexA and RexB may be a factor causing abortive infection. In order for λ to enter the lytic cycle it must avoid the effects of the RexAB system. It is thought that λ is able to upregulate RexB levels thus preventing abortive infection by RexA (Parma et al., 1992).

Certain *E. coli* K12 strains carry the defective lambdoid prophage e14 that carries another abortive infection system (Brody and Hill, 1988; Hill et al., 1989). This Abi system Lit (late inhibitor of T4) is a metallopeptidase that is constitutively expressed in an inactivated form. It is activated by a short peptide sequence called GoI that is located within the N terminal region of the major head protein (gp23) of phage T4 (Bergsland et al., 1990; Copeland and Kleanthous, 2005). When activated, Lit cleaves a conserved motif in the translation elongation factor, Tu, that leads to the inhibition of translation and phage exclusion (Yu and Snyder, 1994).

The Lit system is not the only example of a specific peptide sequence activating an abortive infection system. The clinical *E. coli* strain CT196, carries the *prr* locus which contains the *prrC* gene (Depew and Cozzarelli, 1974). During T4 infection, the phage produces an endonuclease inhibitor, Stp, which inhibits an RM system by binding to *Eco*prrI. However, *Eco*prrI forms a complex with PrrC which is released due to the interaction with Stp (Penner et al., 1995). PrrC is a ribonuclease that cleaves the host $tRNA_{Lys}$ leading to the inhibition of translation and subsequent phage exclusion. Wild type T4 however, is not aborted by PrrC as it is able to repair the damage caused by it. This locus was discovered in T4 mutants that were defective in RNA repair i.e. polynucleotide kinase and RNA ligase mutants (Depew and Cozzarelli, 1974). The Stp protein is not the only activator of PrrC. Several pyrimidine nucleotides, in particular

dTTP is able to directly activate PrrC, and they suggests that there may be functions of PrrC other than acting as an abortive infection system (Amitsur et al., 2003).

Abortive infection systems are not limited to Gram negative bacteria; in fact, the majority of known abortive infection systems have been found in the Gram positive, *Lactococcus lactis*. This is partly due to the fact that *L. lactis* is an important bacterium in the dairy industry. Fermentation cultures of *L. lactis* represent the largest man made bacterial cultures and they are highly susceptible to phage infection (Chopin et al., 2005; Garneau and Moineau, 2011). There are at least 23 identified *L. lactis* Abi systems identified, the abundance of which is in part due to the artificial selection of phage resistance phenotypes during selection of strains for the fermentation process (Chopin et al., 2005; Dy et al., 2014a). *L. lactis* Abi systems (are labelled as Abi followed by a letter) have a wide range of activities from preventing DNA replication of phages, causing the premature lysis of infected cells or interfering with phage packaging of DNA. A list of identified *L. lactis* Abi systems and their modes of actions (if known) is reviewed in (Ainsworth et al., 2014).

The best characterised *L. lactis* Abi system is the AbiD1 system which, like the Lit and PrrC systems of *E. coli*, is also activated by a single phage-encoded protein. Prior to phage infection *abiD1* mRNA is constitutively produced. This mRNA is highly unstable and due to a stem loop structure, the translation of it is also highly inefficient. The stem loop structure is located in the translation initiation region which is able to act as a binding site for phage bIL66 protein ORF1. Binding of ORF1 during phage infection stabilises the mRNA leading to the translation of the transcript (Bidnenko et al., 2002, 2009). The AbiD1 protein then interferes with the function of another phage protein, ORF3, which is a RuvC-like endonuclease (Bidnenko et al., 1998). The role of ORF3 is to resolve branched DNA structures and thus inhibition of ORF3 disrupts phage DNA replication and packaging (Bidnenko et al., 1998). The AbiD1 system can also be activated by a decrease in temperature, implying again that it too is more than “just” an Abi system and could have alternative functions (Bidnenko et al., 2009).

Prior work in this lab identified a cryptic plasmid in the phytopathogen, *P. atrosepticum* SCRI1039 that encodes a protein, $ToxN_{Pa}$ that shared homology with the AbiQ protein of *L. lactis* (Fineran et al., 2009). Subsequent work on $ToxN_{Pa}$ revealed that it belonged to a new Type of toxin-antitoxin family and, like AbiQ, is also capable of aborting

infection by certain phages (Fineran et al., 2009). Further details on toxin-antitoxin systems and their roles in phage resistance are discussed below.

### 1.3 Toxin-antitoxins

Toxin-antitoxin (TA) systems are genetic loci that classically contain two components; a toxin protein and an antitoxin the neutralises the effect of the toxin. These loci were initially discovered as a plasmid maintenance system that worked by post segregationally killing (Gerdes et al., 1986a). A classical toxin-antitoxin system consists of two genes, an antitoxin gene and a toxin gene located downstream under the control of a single promoter. However, this is not always the case as three component toxin-antitoxin systems have been reported (Hallez et al., 2010). There are also examples of TA systems where the toxin gene is located upstream of the antitoxin (Jurenaite et al., 2013). The encoded toxin is always proteinaceous in nature and has a detrimental effect on the cell. The antitoxin component on the other hand can either be a protein or RNA and may or may not directly interact with the toxin. The nature and mechanism of action of the antitoxin forms the basis for the classification of TA systems into their different Types. To date there have been up to six reported Types of TA systems (Harms et al., 2018; Page and Peti, 2016). TA systems are widespread and although originally discovered as a plasmid maintenance system, have been found on bacterial chromosomes as well as phage genomes (Bast et al., 2008; Lehnherr et al., 1993). They are also often found in multiple copies, for example *E. coli* MG1655 contains 19 Type I, 13 Type II and three Type IV TA systems (Harms et al., 2018). Another example, *Mycobacterium tuberculosis* is predicted to contain 88 TA systems (Ramage et al., 2009). It has also been reported that pathogens have a high number of TA systems compared to commensals suggesting a possible role in virulence (Georgiades and Raoult, 2011).

There have been three well established functions of TA systems, namely plasmid maintenance, persister formation and abortive infection (Harms et al., 2018). Plasmid maintenance, the original assigned function of toxin-antitoxin systems relies on the differential stability of the toxin and antitoxin. In general, the antitoxin is less stable than the toxin component. Thus, when a bacterial cell carrying a plasmid with a TA system divides into two, if a daughter cell does not inherit the plasmid it eventually dies due to

the activity of the cytoplasmic toxin. This is because the plasmid is no longer able to provide the antitoxin, thus once any remaining antitoxin is degraded, the more stable toxin would no longer be neutralised. In this way the cells become “addicted” to the plasmid as they are not able to survive without it (Yarmolinsky, 1995). It is thought that chromosomal TA systems may help prevent this addiction by providing their own antitoxins (Bast et al., 2008).

Persister formation is one of the key reasons why TA systems has gathered interest (Lewis, 2010). The phenomenon of persisters was discovered by Joseph Bigger, who found that, after treating *S. aureus* with penicillin, a small percentage of cells survived. However after growing and exposing the survivors to penicillin again, the percentage of survivors was similar in both cases (Bigger, 1944). He concluded that these survivors were not antibiotic resistant mutants but rather antibiotic tolerant. As this phenotype was transient it was not due to a mutation but rather a mechanism that causes a small fraction of the bacterial population to enter a dormant non-replicating state, called persistence. After the removal of the antibiotic, persisters are able to return to a normal replicative state and become sensitive to antibiotics again. The role of TA systems and persister formation has become a controversial one recently. It is clear that TA systems are linked to the generation of persisters (Dörr et al., 2010; Tripathi et al., 2014). However, there have been many misconceptions in the field that have confused any definition of the exact role of TA systems with regards to persister formation (Kim and Wood, 2016). Furthermore, several high profile papers from one research group have been retracted recently, apparently due to data being affected by phage contamination (Harms et al., 2017). Originally that group found that 10 TA systems in *E. coli* were involved in persister formation, but it seems that this result is not repeatable (Harms et al., 2017; Maisonneuve et al., 2013, 2018). Although controversial, it had been commonly thought that the stochastic production of ppGpp, and subsequently the lon protease, are able to generate persisters by degradation of the antitoxin (Maisonneuve et al., 2013). It has been shown that persistence is unaffected in cells are unable to produce ppGpp or the lon protease (Chowdhury et al., 2016; Shan et al., 2015). With the recent retraction of the Gerdes group papers this is now no longer thought to be the case and so the mechanism by which TA systems are activated to generate persisters is now unclear.

Several TA systems from various Types confer phage resistance (Harms et al., 2018). Whether this is a result of “altruistic suicide” or halting of bacterial growth caused by TA systems is still to be determined (Song and Wood, 2018). Abortive infection due to TA systems is a main focus of this thesis and this will be discussed in further detail later.

Other than the three well established roles of TA systems, other functions of TA systems have been reported. The MqsR/MqsA system appears to have several functions. MqsR, the toxin, acts as a specific RNase targeting the sequence GCU independently of ribosomes, leading to changes in cellular metabolism affecting quorum sensing and biofilm formation (Yamaguchi et al., 2009). Furthermore, MqsR is able to activate another TA system, GhoT/GhoS by cleaving the mRNA of the antitoxin, GhoS leading to increased persistence (Cheng et al., 2014; Wang et al., 2013). The antitoxin component, MqsA, also plays a role other than neutralising MqsR activity. MqsA is able to repress RpoS, the master regulator of stress, through binding to the RpoS promoter. The repression of *rpoS* leads to the reduction of cyclic diguanylic acid and subsequently a decrease of biofilm formation and an increase in cell motility (Wang et al., 2011). MqsA can also reduce biofilm formation by binding to the promoter of the biofilm regulator, CsgD (Soo and Wood, 2013). The MqsR/MqsA system is not the only TA system involved in regulating RpoS. The YafQ/DinJ system’s antitoxin (DinJ) can indirectly repress RpoS by inhibiting CspE a positive regulator of RpoS (Hu et al., 2012). The roles of TA systems also extend to virulence as the *S. aureus* virulence genes, *hla* and *efb* can be repressed by the antitoxin SavR and in complex with its toxin, SavRS, by binding to the promoter regions of both *hla* and *efb* (Wen et al., 2018). TA systems have been assigned to different Types based on the nature and mode of neutralisation of their antitoxins (**Figure 1.3**). The details of each type and their involvement in phage resistance are discussed in the ensuing sections.

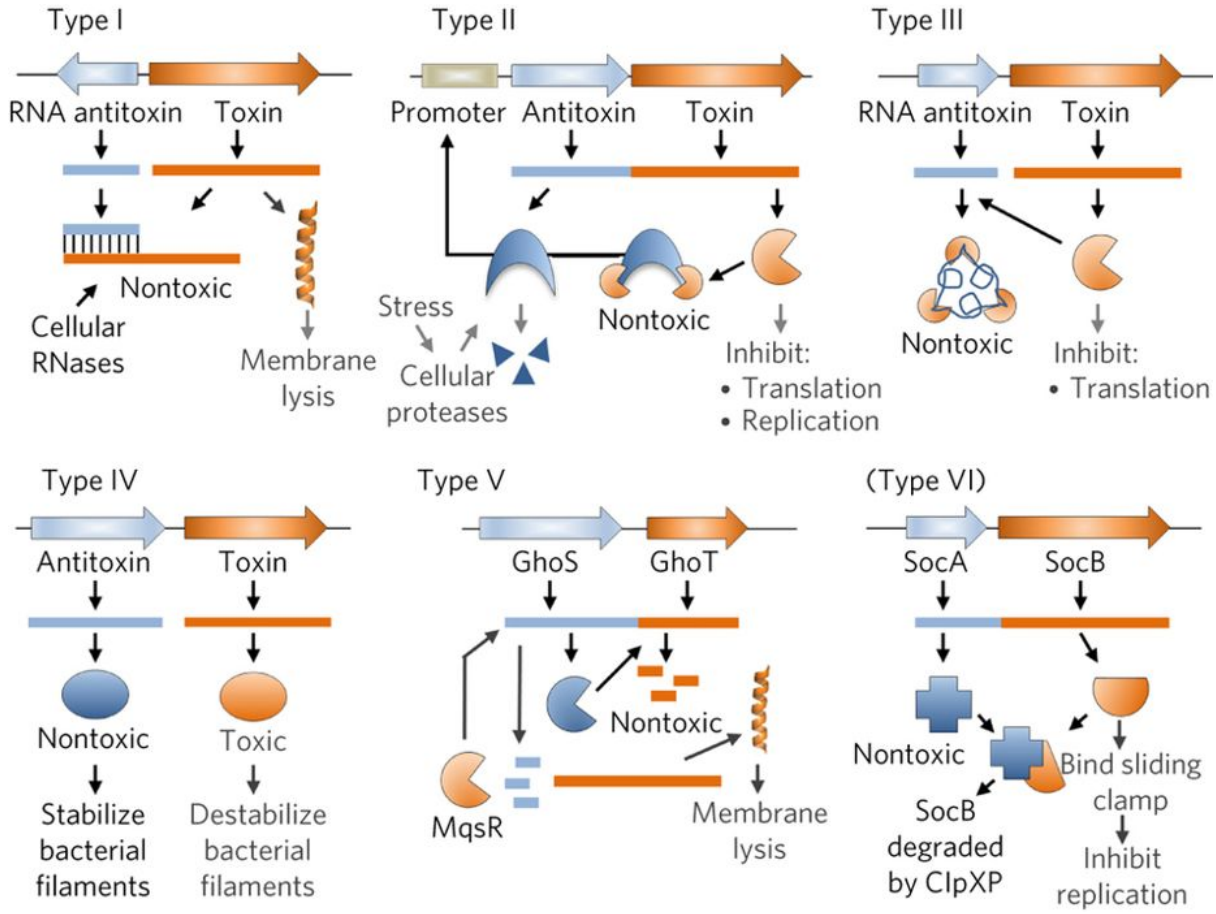



**Figure 1.3** The six Types of TA systems

Toxins are shown in orange with antitoxins in blue. Only Type I and III antitoxins are RNA; all other antitoxins are proteins. A Type VII TA system has been described (not shown here), with further details provided on its subheading. This figure has been taken from (Page and Peti, 2016).

### 1.3.1 Type I

The original TA system and the founding member of the Type I TA systems is the Hok-Sok system. It was discovered on plasmid R1 where it was shown to be involved in the maintenance of the plasmid (Gerdes et al., 1986b). Type I TA systems have been found in many bacterial species (Fozo et al., 2010). They are defined as having an antitoxin that is the antisense RNA sequence of the toxin mRNA counterpart. The antitoxin of Type I TA systems form a duplex with the mRNA of their toxins that is targeted by RNAse III, hence neutralisation is achieved by preventing the translation into the toxic protein (Gerdes et al., 1986b; Thisted et al., 1994). The toxin component of Type I TA systems are mostly small membrane-associated proteins which function as phage holins by forming pores in the bacterial membrane disrupting the proton motive force (Brantl and Jahn, 2015). Type I toxins such as SymE and RalR can also function as RNases or DNases, respectively (Guo et al., 2014; Kawano et al., 2007).

With regards to phage resistance, the Hok-Sok system has been shown to exclude phage T4, albeit only very modestly. There was only a 42% reduction in efficiency of plating in the presence of the *hok*/*sok* locus within a medium copy (pBR322) plasmid vector (Pecota and Wood, 1996). The activation mechanism of the Hok-Sok system still remains unknown, however, it has been suggested that the decrease of host translation caused by the phage and subsequent decrease of the antitoxin, Sok is the cause (Pecota and Wood, 1996). This may explain why phage lambda is not excluded by the system as host transcriptional machinery is not shut down during its infection. However, it still does not explain why other phages are not excluded. The authors suggest that there may be other unknown mechanisms developed by these phages to prevent the activation of the toxin (Pecota and Wood, 1996). No other Type I TA system has been reported to confer phage resistance, thus far.

### 1.3.2 Type II

The Type II TA systems are the best characterised TA systems, reflecting a wide research interest in the formation of persisters, although their role in persister formation might not be as important as previously thought (Goormaghtigh et al., 2018). Like Type I TA systems, these are found in almost all bacterial genomes and are also present in phage genomes (Leplae et al., 2011; Pandey and Gerdes, 2005). The Type II TA

system is defined as having a protein antitoxin that directly interacts with the toxic protein. The formation of the toxin-antitoxin protein complex causes the neutralisation of the toxin. The antitoxin and toxin are under the control of a single promoter. The antitoxin is also less stable than its toxin counterpart, which is partly due to it being partial folded, thus making it more susceptible to degradation by proteases (Cherny et al., 2005). The Type II TA toxins are released upon the degradation of their antitoxins when the stringent response activates the Clp or Lon protease (Brzozowska and Zielenkiewicz, 2013).

The regulation of Type II TA systems is more complex than just the binding of an antitoxin to a toxin. The antitoxin protein consists of two domains, a C-terminal domain that is used for the protein-protein interaction with the toxin protein and an N-terminal domain containing a DNA-binding domain. There are three different DNA-binding domains an antitoxin may have depending on how they fold (Chan et al., 2016). The role of these DNA-binding domains is to provide a further level of regulation as they allow the antitoxin to bind to the operator region of their own operons. Without the binding of the toxin the antitoxin binds weakly to the operator, however the toxin, when in complex with the antitoxin, changes conformation allowing it to bind more strongly and repress transcription (Magnuson et al., 1996; Overgaard et al., 2009). The toxins and antitoxins can also form different complexes with each other, with different affinities for the operator to ensure an optimal ratio of toxin-to-antitoxin in a process called conditional cooperativity (Cataudella et al., 2012). For example in the RelBE system, an abundance of antitoxin, RelB, leads to the formation of a $RelB_2RelE$ heterotrimer consisting of 2 RelB antitoxins and a single RelE toxin (Overgaard et al., 2009). This complex binds strongly to stop transcription of the TA operon. However, should the antitoxin levels decrease, possibly in response to cellular stress, the $RelB_2RelE$ complexes are disrupted by free toxin and there is a higher formation of $RelB_2RelE_2$ complexes which are unable to bind the promoter (Overgaard et al., 2009). This prevents the $RelB_2RelE$ complex from repressing the operon, thus more antitoxin is translated.

The toxins and antitoxins of Type II TA systems are grouped into families based on amino acid similarity (Pandey and Gerdes, 2005; Van Melderen and Saavedra De Bast, 2009). There are currently 12 families of toxins and 13 families of antitoxins (Makarova et al., 2009). The Type II TA systems appear to be able to use different

families of antitoxins to neutralise the same family of toxins and *vice versa* for the toxins (Guglielmini and Van Melderen, 2011). The toxins of Type II TA systems have a wide range of activities. This includes inhibition of DNA gyrase, acting as an endoribonuclease and inactivating elongation factors (Bernard et al., 1993; Christensen-Dalsgaard et al., 2010; Cruz et al., 2014; Zhang et al., 2003).

Type II TA systems have been linked to phage resistance. For instance, the deletion of the chromosomally-located *mazEF* locus results in a more efficient P1 infection (Hazan and Engelberg-Kulka, 2004). This highlights that chromosomal TA systems may function as effective anti-phage systems. Two homologous Type II TA systems, RnlAB and LsoAB, have also been shown to abort T4 mutants defective in the Dmd protein (Otsuka and Yonesaki, 2012). It has been shown that, during a T4 infection, the levels of the antitoxins from these two systems drop dramatically (Koga et al., 2011; Otsuka and Yonesaki, 2012). In wild type T4 the Dmd protein prevents abortive infection by acting as an antitoxin to both toxins (Otsuka and Yonesaki, 2012). These examples highlight again that phages are capable of evolving specific solutions to prevent abortive infection, but also show that the levels of antitoxin play a key role in TA activation, further emphasised in Type II TA systems by the complexity of their regulation.

### 1.3.3 Type IV

Like Type II, the Type IV TA system also has a proteinaceous antitoxin, however it does not directly interact with the toxin protein, a hallmark of a Type IV TA system. Instead the antitoxin neutralises the effects of the toxin by promoting the opposite effect caused by the toxin. For instance, in the CbeA-CbtA system the toxin CbtA (originally called YeeV) binds and inhibits MreB and FtsZ which are required for cell division (Tan et al., 2011). The antitoxin, CbeA, acts as an antagonist and instead promotes the polymerisation of both MreB and FtsZ (Masuda et al., 2012). CbeA can also prevent the toxicity of other MreB and FtsZ inhibitors, thus it is likely that it has a much wider role than just neutralising CbtA (Masuda et al., 2012).

AbiE is a Type IV TA system firstly identified on a plasmid originating from *L. lactis* for its anti-phage activity (Garvey et al., 1995). The AbiE system consists of the AbiEi antitoxin and the AbiEii toxin which do not interact with each other (Dy et al., 2014b).

The exact mechanism by which AbiEii causes toxicity is unknown but is linked to its ability to bind GTP and its predicted nucleotidyltransferase activity (Dy et al., 2014b). The antitoxin, AbiEii, which is under control of the same promoter, contains a DNA binding domain. Similar to Type II TA systems, the antitoxin negatively autoregulates the transcription of the operon by repressing the AbiE promoter (Hampton et al., 2018). However, in contrast to the Type II TA systems, the AbiEi and AbiEii, protein were shown not to interact with each other, based on the results of coimmunoprecipitation experiments (Dy et al., 2014b). The AbiE system has been shown to abort *L. lactis* phage 712, by an unknown mechanism. However as 712 DNA replication is unaffected during infection of a host carrying the AbiE system, it is thought to affect phage packaging or release (Garvey et al., 1995). The toxin AbiEii is toxic even in the absence of phage so it is likely that the toxic effect on the bacterial cell itself causes abortive infection (Dy et al., 2014b). It is not known what component of the phage activates the AbiE system.

### 1.3.4 Type V

The GhoST system is the only Type V TA system to date and is defined as a protein antitoxin that neutralises the effects of the toxin by directly degrading the toxin mRNA. The antitoxin of the GhoST system, GhoS, acts as an RNAse that is able to cleave the mRNA of GhoT, the toxin (Wang et al., 2012). This TA system was named after the effects that the toxin causes, as GhoT is a small hydrophobic membrane associated peptide that disrupts the cell membrane causing the formation of ghost cells (Wang et al., 2012).

Phage resistance has not yet been examined in the GhoST system. However, the GhoST is regulated by the Type II system, MqsR/MqsA. Although MqsR/MqsA has not been associated with phage resistance, other Type II systems have been (Cheng et al., 2014; Hazan and Engelberg-Kulka, 2004; Otsuka and Yonesaki, 2012; Wang et al., 2013).

### 1.3.5 Type VI

Like the Type V system, there is also only one currentle example of a Type VI system. It is defined as having an antitoxin that interacts with its toxin component. Neutralisation however, does not occur by this interaction but by the promotion of degradation caused by this binding (Aakre et al., 2013). The only Type VI TA system, SocAB, consists of the SocA antitoxin and the SocB toxin. While both proteins are constitutively expressed, the antitoxin, SocA, acts as a proteolytic adaptor, that binds to SocB and targets it for degradation by the ClpXP protease (Aakre et al., 2013). Toxicity of SocB is caused by its binding to the β sliding clamp DnaN, inhibiting replication elongation (Aakre et al., 2013). SocB is predicted to bind to a hydrophobic cleft that is also the binding site of several other proteins important for DNA repair and replication, such as Pol IV, Pol V, MutS, MutL and Hda (Robinson et al., 2010). The interaction of the SocAB system has not yet been investigated with regards to phage infection.

### 1.3.6 Type VII

A seventh Type of TA system has been described where the neutralisation of the toxin is due to the antitoxin acting as an enzyme, causing the oxidation of a cysteine residue on the toxin (Marimon et al., 2016). The Hha/TomB TA system is an example of a Type VII TA system where the toxic Hha protein can be neutralised upon oxidation by the antitoxin, TomB. Like other TA systems the Hha/TomB TA system genes are contiguous and operonic. It was originally discovered as a regulator of biofilm and motility (Barrios et al., 2006). The Hha toxin affects multiple functions of the cell. It can interact with the H-NS protein involved in bacterial motility control (Nieto et al., 2002). Hha can also bind to tRNA codons and induce the activation of prophage lytic genes, leading to cell death (García-Contreras et al., 2008). The overexpression of Hha can also induce expression of ClpP, ClpX and the lon protease, all of which are involved in Type II TA activation, which may contribute to cell lysis (García-Contreras et al., 2008).

Phage resistance via abortive infection has not be reported in the Hha/TomB TA system. However, Hha has been shown to upregulate OmpA (Balsalobre et al., 1999). OmpA is an important receptor for multiple phages and so it is possible that the Hha/TomB TA system may affect phage replication by other means (Morona et al.,

1984). Furthermore, the protein product of gene 5.5 of phage T7 binds and inhibits H-NS, showing another example of phage interaction with this system, although indirect (Liu and C Richardson, 1993).

### 1.3.7 Type III

#### 1.3.7.1 Definition and genetic organisation

The Type III TA systems are defined as having a non-coding RNA antitoxin that physically interacts with the toxic protein forming a complex to neutralise it (Fineran et al., 2009). Similar to other TA systems, Type III TA systems are bicistronic containing an antitoxin gene preceding a toxin gene under the control of a single promoter (Fineran et al., 2009) (**Figure 1.4**). The antitoxin carries a series of short repeating sequences that, as RNA, are cleaved into single subunits by the toxin protein (Bélanger and Moineau, 2015; Fineran et al., 2009). Additionally, the antitoxin and toxin components are separated by a Rho-independent terminator which can reduce transcription of the full transcript by up to 90% (Fineran et al., 2009). The multiple antitoxin copies and the terminator in a Type III TA locus ensure that more antitoxin is produced than toxin. Like the other classical TA systems i.e. Type I and Type II, the Type III TA systems are widespread with over 600 putative systems identified bioinformatically, encompassing 8 phyla (Goeders et al., 2016).

The Type III TA systems can be divided into three families based on sequence homology of their toxin components. The three families, ToxIN, TenpIN, CptIN, follow a naming system where the I and N letters refer to the antitoxin and toxin component respectively (Blower et al., 2012b) (**Figure 1.4**). For instance, ToxIN refers to the entire system whereas ToxI refers specifically to the antitoxin, likewise ToxN for the toxin. For the ToxIN system of *Pectobacterium atrosepticum*, the first identified Type III TA system, the letters Pa in lower case after ToxIN is used i.e. $ToxIN_{Pa}$ (Blower et al., 2012b). The TenpIN family is named after the *Photorhabdus* Type III TA system, and is abbreviated from **T**ype III **En**dogenous to ***P****hotorhabdus* **I**nhibitor/toxi**n** (Blower et al., 2012b). Again, the TenpIN system from *Photorhabdus luminescens* is specifically referred to as $TenpIN_{Pl}$. The final family, CptIN, is abbreviated from *Coprococcus* Type III Inhibitor/toxin, after the GD/7 system of *Coprococcus catus* (Blower et al., 2012b).

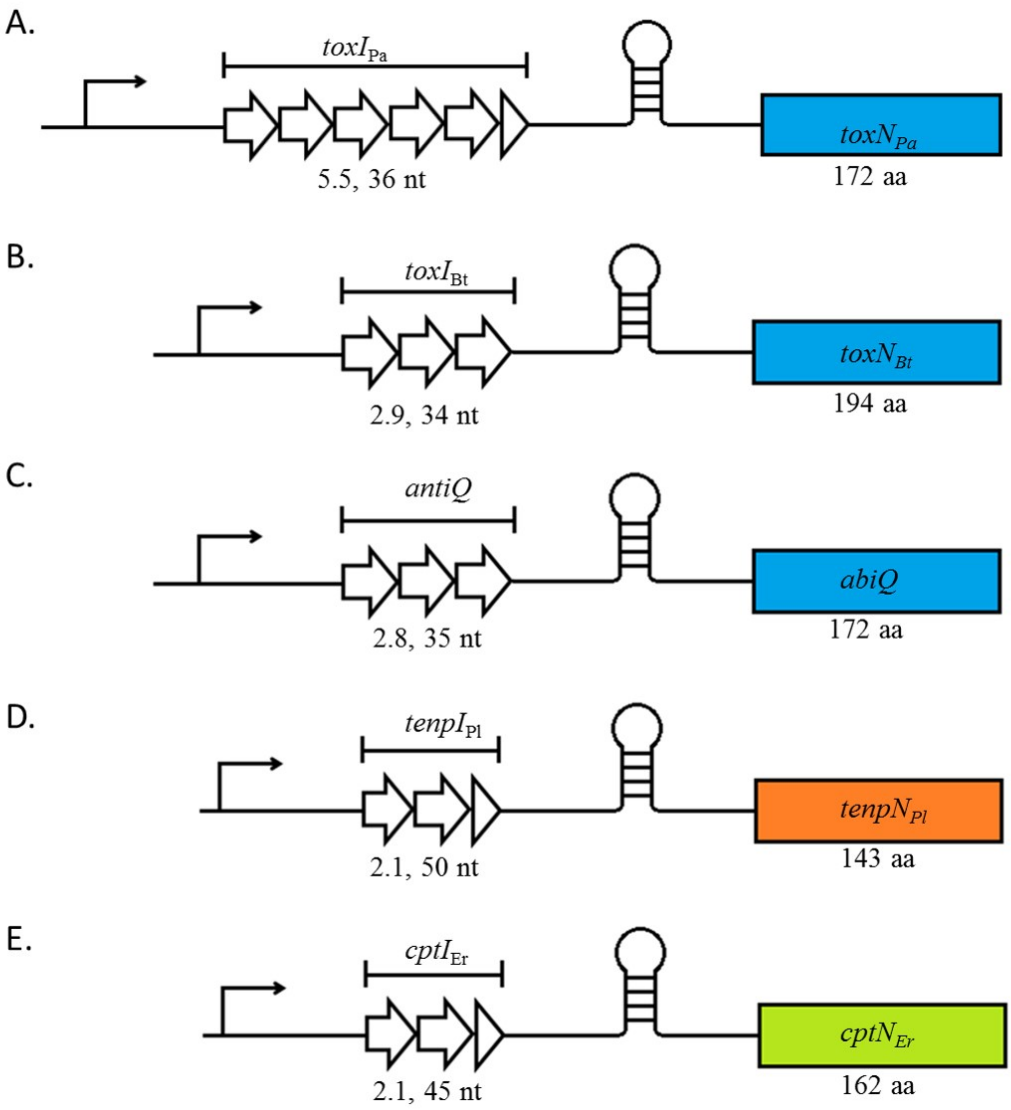



**Figure 1.4** The genetic organisation of the three families of Type III TA systems

The TA locus is under control of a single promoter that is immediately followed by a sequence of repeats, that is the antitoxin, which is separated from the toxin by a rho independent terminator. **A-C** shows three TA systems belonging to the ToxIN family (toxin genes highlighted in blue) For the AbiQ system, the toxin gene is called *abiQ* and the antitoxin gene called *antiQ*. **D** shows the locus of the $TenpIN_{Pl}$ system with the toxin gene, $TenpN_{Pl}$ in an orange box. **E** shows an example of the CptIN family, $CptIN_{Er}$, with the gene for the toxin component $CptN_{Er}$ as a green box. This figure was taken from (Goeders et al., 2016) and made by the author of this study.

While the Type III TA systems are split into their three families based on their toxins, their antitoxin sequence appears to be less conserved. Three characterised ToxIN family members, $ToxIN_{Pa}$, $ToxIN_{Bt}$ and the AbiQ system do not share any antitoxin sequence homology, although the lengths of their antitoxins are somewhat conserved. Interestingly the number of repeats of each antitoxin between Type III TA systems also differ. The number of repeats have been shown to be important in neutralising the toxin component. For the $ToxIN_{Pa}$ system, which naturally has 5.5 repeats, only 2.5 repeats were crucial for toxin inhibition (Blower et al., 2009). This is different for the AbiQ system where only 1.8 antiQ repeats were necessary (Bélanger and Moineau, 2015). Additionally, it has been shown that the anti-phage capability of AbiQ is affected by the number of repeats (Bélanger and Moineau, 2015).

#### 1.3.7.2 Structures of Type III TA systems

The Type III TA systems form complexes held together by RNA-protein interactions. The structures of three of these complexes have been solved (Blower et al., 2012a; Rao et al., 2015; Short et al., 2013). The two ToxIN family members, $ToxIN_{Pa}$ and $ToxIN_{Bt}$ form trimeric heterohexamers where three toxin proteins are located at the apices of the triangles which are held together by three antitoxin subunits (**Figure 1.5**). All characterised Type III TA system toxins have been identified as endoribonucleases and recognise a short specific sequence that is found throughout the bacterial genome but also within the antitoxin repeat sequence. The specific sequence recognised is different between different Type III toxins. The toxin cleaves the antitoxin at this location into their single subunits and forms the complexes. In vitro the $ToxN_{Pa}$ toxin has been shown to be able to self-assembly into the complexes using either the non-cleaved $ToxI_{Pa}$ or cleaved $ToxI_{Pa}$ suggesting that complex formation is spontaneous and does not require any other cellular factors (Short et al., 2013). However, the preferred substrate of $ToxN_{Pa}$ is non cleaved $ToxI_{Pa}$ suggesting that cleavage of antitoxin itself is part of the neutralisation process (Short et al., 2013). Recently, it has also been shown that this complex is perhaps not as stable as previously thought and the constant cleavage of newly transcribed antitoxin by the toxin is the main mechanism of neutralisation (Short et al., 2018).

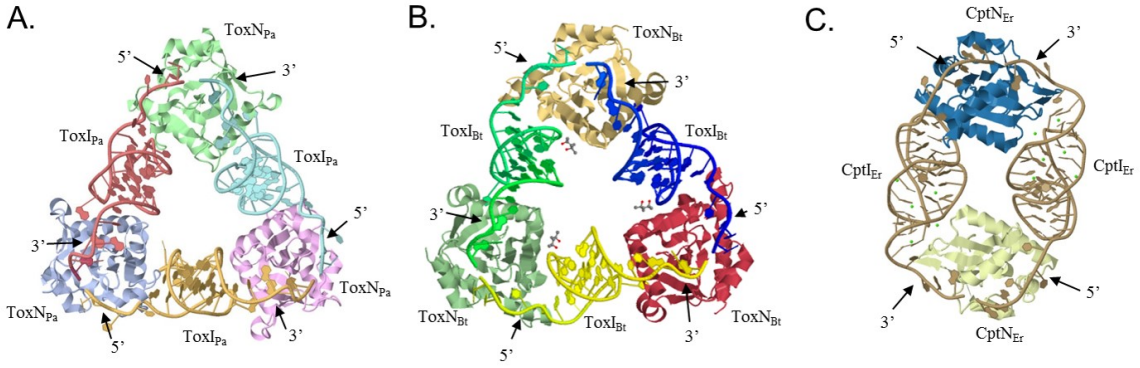



**Figure 1.5** The Type III TA systems form RNA-protein complexes

**A**. The trimeric heterohexamer of ToxIN$_{Pa}$ consists of three toxin proteins and three RNA pseudoknots. **B**. The ToxIN$_{Bt}$ system forms a similar complex to the ToxIN$_{Pa}$ system. **C**. The CptIN$_{Er}$ complex is a heterotetramer consisting of two toxic proteins held together by two RNA antitoxins. This figure was taken from (Goeders et al., 2016) and made by the author of this study.

The ToxIN family members whose antitoxins have been solved, $ToxI_{Pa}$ and $ToxI_{Bt}$, have been shown to fold into a classic H-type pseudoknot containing two single stranded tails at either end of the structure (Blower et al., 2011; Short et al., 2013). These are common RNA structures characterised by three different sections of the RNA strand, looped around each other and held together by a central core consisting of duplex and/or triplex hydrogen bonds deriving from base-pair interactions. For both $ToxI_{Pa}$ and $ToxI_{Bt}$, the core consists of three triplexes in which the third triplex formed of GUU has an interdigitation of a guanine that acts to separate the two stems of the knot (**Figure 1.6A-B**). Given the low sequence homology between $ToxI_{Pa}$ and $ToxI_{Bt}$, it was unexpected that such a similar pseudoknot structure would be formed. However, the antitoxins are unable to complement each other and specificity is achieved by the specific interactions each ToxN makes with its own ToxI. Each ToxN contains electropositive grooving in which the protein residues and RNA bases form hydrogen bonds. The tails at either end of the pseudoknot play a key role in this. The active site of $ToxN_{Pa}$ consisting of Tyr29, Lys33, Thr52, Ser53 and Lys55 interacts and hold the 3' end of the ToxI tail in place (Blower et al., 2011). For the $ToxIN_{Bt}$ system, a single $ToxN_{Bt}$ interacts with the 5' tail of the pseudoknot. The same $ToxN_{Bt}$ also interacts via Lys148 with C19 and G20 from triplex one (Short et al., 2013). A different $ToxN_{Bt}$ then interacts with the 3' tail as well the U10 which binds to a hydrophobic pocket located on the same $ToxN_{Bt}$ (Short et al., 2013). The difference in the ToxN-ToxI interactions between the two TA systems is a result of the differences in sequence specificity. $ToxN_{Pa}$ specifically cleaves RNA at AA✂AU/G whereas $ToxN_{Bt}$ cleaves A✂AAAA (Short et al., 2013).

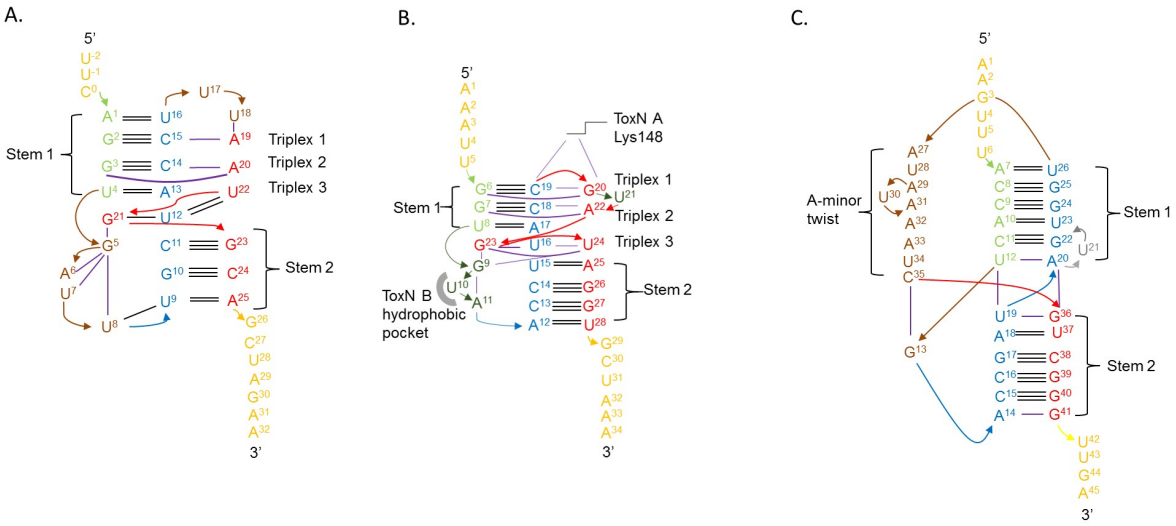



**Figure 1.6** The pseudoknot structures of Type III antitoxins

**A**. The $ToxI_{Pa}$ antitoxin folds into a H type pseudoknot containing a triple triplex core. **B**. The $ToxI_{Bt}$ antitoxin exhibits a similar fold to $ToxI_{Pa}$. **C**. $CptI_{Er}$ has a different fold containing a rare A-minor twist structure. The colours of the bases represent the corresponding base in each of the three antitoxins. Black lines represent hydrogen bonds between bases. This figure has been taken from (Goeders et al., 2016) and made by the author of this study.

The third Type III TA structure to be solved as a complex belongs to the CptIN family. The $CptIN_{Er}$ complex differs from the two ToxIN family complexes as it instead forms a heterotetramer consisting of two toxins and two antitoxins (**Figure 1.5C**). The reason why $CptIN_{Er}$ forms a different complex may be the result of the difference in length of its antitoxin $CptI_{Er}$ (**Figure 1.4E**). The $CptI_{Er}$ is longer than the ToxI antitoxins and also has a different pseudoknot fold giving it flexibility to adopt its distinctive complex form (**Figure 1.6C**). While $CptI_{Er}$ also folds into an H-type pseudoknot it does not contain a core containing triplex bonds and forms instead a counter-clockwise A minor twist motif (Rao et al., 2015). There hasn't yet been a structure solved for a Type III TA system belonging to the TenpIN family but for $TenpIN_{Pl}$, given that it has a long antitoxin, it is predicted to form a similar complex to what is seen for $CptIN_{Er}$ (**Figure 1.4D**).

While the ToxI antitoxins exhibit a similar overall fold to each other that differs from that of the CptI example, the toxin proteins also follow a similar pattern. The $ToxN_{Pa}$ and $ToxN_{Bt}$ have very similar folding that includes a six anti-parallel β-sheet and five variable loop core (Blower et al., 2011; Short et al., 2013). The loops play a role in the interaction with the antitoxin, with one of the loops forming part of the active site (Blower et al., 2011; Short et al., 2013). One of the distinctive features of $ToxN_{Pa}$ and $ToxN_{Bt}$ is a kink in helix three (**Figure 1.7A-B**). The AbiQ toxin, which also has a similar fold to both $ToxN_{Pa}$ and $ToxN_{Bt}$ also has this kink (**Figure 1.7C**). The regions corresponding to the active sites are also similar, for instance, the RNAse activity of AbiQ can be abolished by a Ser51Leu mutation. However, substitution of Ser51 with a threonine, the corresponding residue in the active ToxN maintains its activity. In contrast, while the $CptN_{Er}$ toxin also contains a similar structure consisting of anti-parallel B-sheets it does not contain the distinctive kink in helix three (**Figure 1.7E**). Instead the corresponding helix, helix 4, is considerably shorter. It is also known that the Type II toxins belonging to the MazF, Kid and CcdB families also share similar folding (Goeders et al., 2016). While MazF and Kid are endoRNAses, similar to the Type III toxins, CcdB family toxins are DNA gyrase inhibitors and yet have similar β core folds. However, the active sites are non-conserved, even if they do share the same activity, e.g. ToxN and Kid (Goeders et al., 2016). Interestingly, the absence of the kink in the Kid structure, means it more closely resembles the $CptN_{Er}$ toxin (**Figure 1.7D-E**).

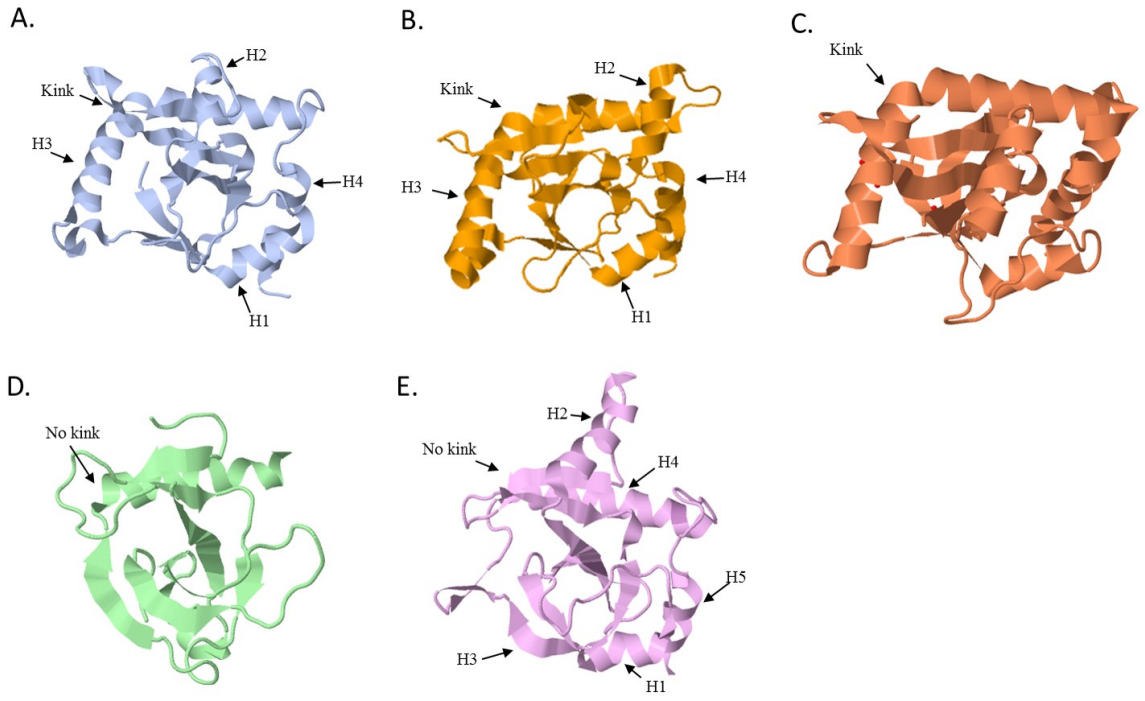



**Figure 1.7** Solved structures of Type III toxins and the Type II toxin, Kid
**A**. $ToxN_{Pa}$, **B**. $ToxN_{Bt}$, **C**. AbiQ, **D**. Kid and **E**. $CptN_{Er}$. All toxins shown are endoRNases and exhibit similar folds. **A**, **B** and **C** contain an obvious kink which is notably absent in **D** and **E** (labelled). This figure has been taken from (Goeders et al., 2016) and made by the author of this study.

#### 1.3.7.3 Functions of Type III TA systems

To date, Type III TA systems have been shown to have at least two roles, plasmid maintenance and phage resistance through abortive infection.

##### 1.3.7.3.1 Plasmid maintenance

As described previously, TA systems were originally identified through their role in plasmid maintenance through post segregationally killing. This relies on an unstable antitoxin and a more stable toxin where loss of the plasmid quickly results in cellular death because the antitoxin is degraded quicker, leading to cell death through the release of the toxin. Thus only cells that have maintained the plasmid are able to survive. The toxins of Type III TA systems have been shown to be bacteriostatic, where a period exists when production of the antitoxin can rescue the cells and allow replication again (Fineran et al., 2009). It is predicted that a point of no return exists, where production of the antitoxin no longer rescues the cell and cellular death soon follows. This point has not been defined precisely but is thought to occur after eight hours for the $ToxIN_{Pa}$ system which suggests that its plasmid maintenance capability is not through post segregationally killing but rather the cessation of growth in TA-plasmid free cells (Fineran et al., 2009).

The $ToxIN_{Pa}$ and the $CptIN_{Er}$ systems have been shown to retain up to 100% of plasmids in an *E. coli* W3110 background compared to just 50% in the controls with no Type III TA system (Rao et al., 2015; Short et al., 2013). While the $ToxIN_{Pa}$ system was originally identified on a plasmid, the $CptIN_{Er}$ system is located on the chromosome of *Eubacterium rectale*. Thus it is curious that it can function as a plasmid maintenance system in *E. coli* (Blower et al., 2012b). In contrast the $ToxIN_{Bt}$ system, which is naturally found on a plasmid had no effect on plasmid retention in *E.coli* compared to the control (Short et al., 2013). The likely reason for this is that $ToxN_{Bt}$ is less toxic in *E. coli* compared to $ToxN_{Pa}$, suggesting that its diminished toxicity on *E. coli* is not strong enough to allow plasmid retention. In contrast the $ToxIN_{Bt}$ system has been able to act as a plasmid maintenance system in its natural host, *Bacillus subtilis*. A part of the *B. subtilis* lifecycle is sporulation, in which the bacteria enter a dormant form, known as a spore. As a spore the bacterium, is highly resistant to hostile environmental conditions, showing similar traits to the persistence phenotype

previously described. Likewise, sporulation is reversible, and the bacteria are able to revert back to an active state when conditions are favourable. During sporulation the rate of plasmid loss can be as high as 95% (Turgeon et al., 2008). The $ToxIN_{Bt}$ system has been shown to reduce plasmid loss during sporulation by a factor of 10 (Short et al., 2015). It was shown that the total number of mature spores decreased as a result of toxin activation in forespores that have lost the plasmid (Short et al., 2015).

#### 1.3.7.3.2 Abortive infection

The first Type III TA system to be identified was $ToxIN_{Pa}$ from a cryptic plasmid originating from *P. atrosepticum* SCRI1039 (Fineran et al., 2009). The $ToxN_{Pa}$ protein showed homology with the abortive infection system AbiQ, of *L. lactis* and likewise was able to abort several *P. atrosepticum* phages (Fineran et al., 2009). The first *P. atrosepticum* phage to be characterised in detail is ΦTE, a transducing *Myoviridae* member (Blower et al., 2012a). ΦTE was strongly aborted by the $ToxIN_{Pa}$ system however, rare escapes could be isolated at a frequency of $10^{-8}$ (Blower et al., 2012a). Characterisation of these insensitive escape phages revealed that they had an expanded region in a non-coding region of their genome that could produce an antitoxin mimic. One ΦTE escape appeared to have undergone recombination with the $ToxI_{Pa}$ region and so was capable of escape by producing the natural $ToxI_{Pa}$ (Blower et al., 2012a). The production of these antitoxin mimics resulted in total insensitivity of these ΦTE escapes to the $ToxIN_{Pa}$ system. No other types of ΦTE escape were identified in this system, and thus the activation mechanism of $ToxIN_{Pa}$ could not be determined.

In the search to identity novel phage escape mutants, the $ToxIN_{Pa}$ system was cloned onto a plasmid and shown to abort phages in heterologous hosts such as *E. coli* and *Serratia* ATCC 39006 (*S*39006) (Fineran et al., 2009). Prior to that study, a large number of coliphages were found to be sensitive to the $ToxIN_{Pa}$ system, but none were able to evolve escapes (unpublished, 2009-2017, Salmond lab). In *S*39006, a group of three highly similar *Tevenvirinae* members, ΦCHI14, ΦX20 and ΦCBH8 were isolated that were aborted by $ToxIN_{Pa}$ but capable of evolving escapes at frequencies ranging from $10^{-6}$ to $10^{-8}$ (Chen et al., 2017). Three different escape loci were identified in these phages. The most common was a large deletion in a region capable of coding for numerous hypothetical proteins, or by mutations in the *asiA* gene or a single escape

that had a mutation in an ORF of unknown function, ORF84 (Chen et al., 2017). As the large deletion region and ORF84 had no obvious predicted functions, their roles in ToxIN$_{Pa}$ activation could not be determined. The *asiA* gene function is a well-known activator of middle gene transcription in phage T4 (Colland et al., 1998a). As a middle gene activator, AsiA is capable of affecting numerous middle genes thus it role in abortive infection is not clear.

The other well studied Type III TA system with regards to abortive infection is the AbiQ system. AbiQ was able to abort a wide range of phages and like the ToxIN$_{Pa}$ system, functions as an anti-phage system when transferred into a heterologous host (Samson et al., 2013). Three different lactococcal phages, P008, bIL170 and Phage c2 were aborted by AbiQ and were able to evolve escapes. Analysis of these escapes revealed mutations in a single protein in all three phages (Samson et al., 2013). These were three different proteins in all three phages but all proteins had unknown functions. Finally, the coliphage, Phage 2, was found to be aborted by *E. coli* carrying the AbiQ system. Escapes of phage 2 were in a putative DNA polymerase. Again, the role that this putative DNA polymerase plays in activating AbiQ is unknown (Samson et al., 2013). While the exact mechanism is not known of AbiQ, there are several key observations that have been made. Firstly, the levels of AbiQ do not increase during phage infection. This has also been observed in the case of the ToxN$_{Pa}$ protein. Secondly, the replication of phage DNA is not prevented during infection in the presence of AbiQ, evident by viral DNA build up. Lastly, while the levels of the AbiQ protein do not decrease, the transcript levels of AbiQ do decrease which may play a role in AbiQ activation (Emond et al., 1998).

One other Type III TA system has been shown to function as an abortive infection system. The TenpIN$_{Pl}$ system, originating from the chromosome of *Photorhabdus luminescens*, when cloned onto a plasmid and transferred to *E. coli* is capable of aborting several coliphages (Blower et al., 2012b). Interestingly, when transferred to *S*39006, the three *Tevenvirinae* family phages that were aborted by ToxIN$_{Pa}$, were also aborted by TenpIN$_{Pl}$. However, no escapes were detected (Chen et al., 2017). This suggests that the activation of ToxIN$_{Pa}$ and TenpIN$_{Pl}$ is different. Type III TA activation specificity has also been observed between ToxIN$_{Pa}$ and AbiQ. For example, phage T4 is not aborted by the ToxIN$_{Pa}$ system but is strongly aborted by AbiQ and unable to evolve escapes (Fineran et al., 2009; Samson et al., 2013).

## 1.4 Summary and aims of this research

The purpose of this study was to further understand the mechanisms of Type III TA system activation during phage infection. As so few phage escape mutants were previously characterised it was decided that, novel phages should be isolated and tested for sensitivity to the Type III TA systems, namely the ToxIN$_{Pa}$ and TenpIN$_{Pl}$ systems. The same rationale was employed from previous studies in which escapes of phages sensitive to Type III TA systems would be sequenced to identify the nature of their mutation and the role they play in the activation of the TA system. Initially work begun solely on *P. atrosepticum* phages as this was the natural host of the ToxIN$_{Pa}$ system and the most "natural" way that phages might interact with this TA system. Serendipitously, various other coliphages and a single *Serratia* phage were added to the study leading to further interesting novel escape loci being identified. Finally, the TenpIN$_{Pl}$ system which has been shown to be an effective abortive infection system has never been shown to abort any phages infecting the host that it derived from, *P. luminescens*, TT01 (Rao, 2014). Furthermore, a TT01 phage has never been characterised in depth, thus the final chapter investigates the nature of these TT01 phages in greater detail.

# Chapter 2 Material and methods

## 2.1 Media, reagents and solutions

A list of all the media used in this study is shown in **Table 2.1**. Antibiotics and other supplements used are listed on **Table 2.2**. All other solutions are shown in **Table 2.3**. All media and solutions were sterilised either by autoclaving at 121ºC for 20 mins or, if not appropriate, filtered sterilised using a 0.22 µM filter (Millipore).

**Table 2.1** Media used in this study

| Medium | Ingredients per litre |
|---|---|
| Lysogeny broth (LB) | 10 g Tryptone<br>5 g Yeast extract<br>5 g NaCl |
| 2x LB | 20 g Tryptone<br>5 g Yeast extract<br>10 g NaCl |
| LB-agar (LBA) | LB<br>15 g Agar |
| Top agar | LB<br>3.5 g Agar |
| Top agarose | LB<br>3.5 g Agarose |
| Marine broth (MB) [Difco 2216] | 5 g Peptone<br>1 g Yeast extract<br>0.1 g Ferric citrate<br>19.45 g Sodium chloride<br>5.9 g Magnesium chloride<br>3.24 g Magnesium sulfate<br>1.8 g Calcium chloride<br>0.55 g Potassium chloride<br>0.16 g Sodium bicarbonate<br>0.08 g Potassium bromide<br>34 mg Strontium chloride<br>22 mg Boric acid<br>4 mg Sodium silicate<br>2.4 mg Sodium fluoride<br>1.6 mg Ammonium nitrate<br>8 mg Disodium phosphate |
| Marine broth agar (MBA) | MB<br>15 g Agar |

**Table 2.2** Antibiotics and supplements

| Chemical/supplement | Stock solution | Working concentration |
|---|---|---|
| Antibiotics | | |
| Ampicillin (Ap) | 1 $g.ml^{-1}$ in $dH_20$, stored at -20ºC | 100 $\mu g.ml^{-1}$ |
| Chloramphenicol (Cm) | 500 $mg.ml^{-1}$ in EtOH, stored at -20ºC | 50 $\mu g.ml^{-1}$ |
| Kanamycin (Km) | 500 $mg.ml^{-1}$ in $dH_20$, stored at -20ºC | 50 $\mu g.ml^{-1}$ |
| Sodium pyruvate | 10% (w/v) in $dH_20$, stored at 4ºC | 0.1% (w/v) |

**Table 2.3** Solutions used in this study

| Solution | Components |
|---|---|
| Phage buffer | 10 mM Tris-HCl at pH 7.4<br>10 mM $MgSO_4$<br>0.01% gelatin |
| DNA work | |
| 50x Tris-acetate EDTA (TAE) buffer (L) | 242 g Tris base<br>100 ml EDTA<br>57.1 ml glacial acetic acid<br>pH 8.0 |
| Agarose gel | 1.5% agarose in 1x TAE<br>500 $ng.ml^{-1}$ ethidium bromide |
| 6x Loading dye, purple | NEB – B7024S |
| Solution A | 10 mM MES<br>50 mM $CaCl_2$<br>10 mM $MnCl_2$ |
| Solution A with glycerol | 10 mM MES<br>50 mM $CaCl_2$<br>10 mM $MnCl_2$<br>15% (w/v) Glycerol |
| *Photorhabdus* wash buffer | 1 mM HEPES<br>5% (w/v) sucrose<br>pH 7.0 |
| 5x HF buffer | NEB – F518L |

## 2.2 Bacterial strains

A summary of all bacterial strains is shown in **Table 2.4**. All bacterial strains were grown in LB apart from two strains. *Photorhabdus luminescens*, TT01 were grown in LB supplemented with 0.1% (w/v) sodium pyruvate and *Vibrio harveyi* were grown in MB. Liquid cultures were grown overnight on a rotary wheel at 30ºC for Pectobacteria, *Serratia* and *Bacillus* strains. All Rhizosphere isolates were also grown at 30ºC. *Staphylococcus aureus*, *Pseudomonas* and *E. coli* strains were grown at 37ºC. For long term storage of strains; 1 ml of overnight was added to 1 ml 50% glycerol and stored at -80ºC.

**Table 2.4** Bacterial strains used in this study

| Strains | Relevant characteristics | Reference or source |
|---|---|---|
| ***Pectobacterium atrosepticum*** | | |
| SCRI1043 | Wild type | (Bell et al., 2004) |
| SCC34 | Uncharacterised LPS mutant of SCRI1043 | (Blower et al., 2012a) |
| TER2 | *flgN*::Tn, ΦTE-resistant, $Km^R$, derivative of Pba 1043 | (Blower et al., 2012a) |
| SCRI1 | Wild type | *SCRI, Ian Toth |
| SCRI124 | Wildtype | *SCRI, Ian Toth |
| SCRI139 | Wild type | *SCRI, Ian Toth |
| SCRI41 | Wild type | *SCRI, Ian Toth |
| SCRI13 | Wild type | *SCRI, Ian Toth |
| SCRI1001 | Wild type | *SCRI, Ian Toth |
| SCRI1034 | Wild type | *SCRI, Ian Toth |
| SCRI1039 | Wild type | *SCRI, Ian Toth |
| SCRI1044 | Wild type | *SCRI, Ian Toth |
| SCRI1055 | Wild type | *SCRI, Ian Toth |
| SCRI1056 | Wild type | *SCRI, Ian Toth |
| SCRI1058 | Wild type | *SCRI, Ian Toth |
| SCRI1064 | Wild type | *SCRI, Ian Toth |

Table continued overleaf

*SCRI – Now called the James Hutton Institute

**Table 2.4 continued** Bacterial strains used in this study

| Strains | Relevant characteristics | Reference or source |
|---|---|---|
| ***Pectobacterium carotovorum*** | | |
| SCRI101 | Wild type | *SCRI, Ian Toth |
| SCRI112 | Wild type | *SCRI, Ian Toth |
| SCRI114 | Wild type | *SCRI, Ian Toth |
| SCRI122 | Wild type | *SCRI, Ian Toth |
| SCRI132 | Wild type | *SCRI, Ian Toth |
| SCRI198 | Wild type | *SCRI, Ian Toth |
| SCRI103 | Wild type | *SCRI, Ian Toth |
| SCRI106 | Wild type | *SCRI, Ian Toth |
| SCRI115 | Wild type | *SCRI, Ian Toth |
| SCRI116 | Wild type | *SCRI, Ian Toth |
| SCRI124 | Wild type | *SCRI, Ian Toth |
| SCRI127 | Wild type | *SCRI, Ian Toth |
| SCRI130 | Wild type | *SCRI, Ian Toth |
| SCRI166 | Wild type | *SCRI, Ian Toth |
| SCRI192 | Wild type | *SCRI, Ian Toth |
| 39048 | Wild type | (Monson et al., 2018) |
| ***E. coli*** | | |
| DH5α | | Invitrogen |
| EPI300 | | Epicentre |
| MG1655 | | (Blattner et al., 1997) |
| ***Vibrio harveyi*** | Wild type | (Bassler et al., 1994) |
| ***Staphylococcus aureus*** | | |
| H | Wild type | (Mayr-Harting, 1955) |
| ***Pseudomonas aeruginosa*** | | |
| PA01 | Wild type, ATCC 15692 | Laboratory collection |
| ***Salmonella* Typhimurium** | | |
| 5383 | Wild type | Laboratory collection |
| ***Chromobacter violaceum*** | | |
| CV026 | Tn5 mutant of ATCC 31532 | (McClean et al., 2018) |

Table continued overleaf

*SCRI – Now called the James Hutton Institute

**Table 2.4 continued.** Bacterial strains used in this study

| Strains | Relevant characteristics | Reference or source |
|---|---|---|
| ***Citrobacter rodentium*** | | |
| ICC169 | Spontaneous $Nal^R$ derivative of ICC168 | Laboratory collection |
| ***Bacillus*** | | |
| *Bacillus thuringiensis* subsp. *kurstaki* HD-1 | Wild type | Laboratory collection |
| HD-73 | Wild type | Laboratory collection |
| *Bacillus subtilis* subsp. *subtilis* JH642 | Wild type | (Smith et al., 2014) |
| ***Dickeya*** | | |
| *D. solani* MK10 | Wild type | Isolated from potato in Israel (Pritchard et al., 2013a) |
| *D. zeae* NCPPB 3532 | Wild type | Isolated from potato in Australia (Pritchard et al., 2013b) |
| *D. chrysanthemi* NCPPB 402 | Wild type | Isolated from chrysanthemum in United States (Pritchard et al., 2013b) |
| *D. dieffenbachiae* NCPPB 2976 | Wild type | Isolated from dieffenbachia in United States (Pritchard et al., 2013b) |
| *D. dianthicola* NCPPB 453 | Wild type | Type strain from *Dianthus* (Pritchard et al., 2013a) |
| *D. paradisiaca* NCPPB 2511 | Wild type | Isolated from plantain in Colombia (Pritchard et al., 2013b) |
| ***Photorhabdus*** | | |
| *Ph. luminescens laumondii* TT01 | Wild type | (Bennett and Clarke, 2005) |
| *Ph. luminescens laumondii* FR2 | *tenpIN*::$Km^R$ | (Rao, 2014) |
| Δ*flgG* mutant | TT01 Δ*flgG* mutant | (Easom and Clarke, 2008) |
| Δ*motAB* mutant | TT01 Δ*motAB* mutant | (Easom and Clarke, 2008) |
| *Ph. luminescens akhurstii* LN2 | Wild type | Gift from David Clarke (Qiu et al., 2014) |

Table continued overleaf

**Table 2.4 continued.** Bacterial strains used in this study

| Strains | Relevant characteristics | Reference or source |
|---|---|---|
| ***Photorhabdus*** | | |
| *Ph. luminescens laumondii* HP88 | Wild type | Gift from David Clarke (Ghazal et al., 2016) |
| *Ph. luminescens akhurstii* P2H | Wild type | Gift from David Clarke |
| *Ph. luminescens akhurstii* K4D | Wild type | Gift from David Clarke |
| *Ph. asymbiotica* ATCC 43950T | Wild type | Gift from David Clarek (Fischer-Le Saux et al., 1999) |
| *Ph. asymbiotica* Kingcliffe | Wild type | Gift from David Clarke |
| *Ph.* spp Mexico | Wild type | Gift from David Clarke |
| *Ph.* spp Epypt | Wild type | Gift from David Clarke |
| | | |
| ***Xenorhabdus* strains** | | |
| *X. feltiae* | Wild type | Gift from David Clarke |
| *X. nematophila* | Wild type | Gift from David Clarke |
| | | |
| ***Serratia marcescens*** | | |
| DB10 | Wild type | (Flyg et al., 1980) |
| 1695 | Clinical isolate | (Aucken and Pitt, 1998) |
| 0006 | Environmental isolate | (Aucken and Pitt, 1998) |
| 1047 | Clinical isolate | (Aucken and Pitt, 1998) |
| 0026 | Environmental isolate | (Aucken and Pitt, 1998) |
| 4028 | Clinical isolate | (Aucken and Pitt, 1998) |
| 3078 | Clinical isolate | (Aucken and Pitt, 1998) |
| 0038 | Environmental isolate | (Aucken and Pitt, 1998) |
| 1024 | Clinical isolate | (Aucken and Pitt, 1998) |
| 3127V | Clinical isolate | (Aucken and Pitt, 1998) |
| 0035 | Environmental isolate | (Aucken and Pitt, 1998) |
| 3888 | Clinical isolate | (Aucken and Pitt, 1998) |
| 0028 | Environmental isolate | (Aucken and Pitt, 1998) |
| 3255 | Clinical isolate | (Aucken and Pitt, 1998) |
| 0040 | Environmental isolate | (Aucken and Pitt, 1998) |
| 2529 | Clinical isolate | (Aucken and Pitt, 1998) |
| 0030 | Environmental isolate | (Aucken and Pitt, 1998) |
| 3127 | Clinical isolate | (Aucken and Pitt, 1998) |
| 3078V | Clinical isolate | (Aucken and Pitt, 1998) |
| ATCC 274 | Clinical isolate | (Regué et al., 1991) |
| S6 | Clinical isolate | (Yang et al., 1990) |
| SM365 | Laboratory strain | (Sapriel et al., 2003) |
| | | |
| **Rhizosphere isolates** | FAME prediction | |
| 10Ep11 | *Proteus vulgaris* | (Berg et al., 2002) |
| 9Ez25 | *Serratia grimesii* | (Berg et al., 2002) |

Table continued overleaf

**Table 2.4 continued.** Bacterial strains used in this study

| Strains | Relevant characteristics | Reference or source |
|---|---|---|
| **Rhizosphere isolates** | FAME prediction | |
| 9Ez29 | *Serratia grimesii/Pantoea agglomerans* | (Berg et al., 2002) |
| 9Ec15 | *Xenorhabdus nematophila* | (Berg et al., 2002) |
| 2Kr27 | *Kluyvera cryocrescens* | (Berg et al., 2002) |
| 4Rc14 | *Enterobacter intermedius/Serratia grimesii* | (Berg et al., 2002) |
| 4Rz1 | *Enterobacter intermedius* | (Berg et al., 2002) |
| 4RZ11 | *Pantoea agglomerans* | (Berg et al., 2002) |
| 9Rz4 | *Pantoea agglomerans/Serratia grimesii* | (Berg et al., 2002) |
| 3Rc14 | *Pantoea agglomerans/Klebsiella pnuemoniae* | (Berg et al., 2002) |
| 4Rr2 | *Proteus vulgaris/Serratia grimesii/Pantoea agglormerans* | (Berg et al., 2002) |
| 4Rr2 | *Proteus vulgaris* | (Berg et al., 2002) |
| 3Rc3 | *Serratia fonticola* | (Berg et al., 2002) |
| 4Rz6 | *Serratia grimesii/Pantoea agglomerans* | (Berg et al., 2002) |
| 9Rz10 | *Serratia grimesii* | (Berg et al., 2002) |
| 3Rp10 | *Serratia grimesii/Pantoea agglomerans* | (Berg et al., 2002) |
| 4Rx13 | *Serratia odorifera/Pantoea agglomerans* | (Berg et al., 2002) |
| 3Rr8 | *Serratia plymuthica* | (Berg et al., 2002) |
| 4Rc6 | *Serratia plymuthica* | (Berg et al., 2002) |
| 4Rx5 | *Serratia plymuthica* | (Berg et al., 2002) |
| 3Rr9 | *Serratia proteomaculans* | (Berg et al., 2002) |
| 3Rc15 | *Serratia proteomaculans* | (Berg et al., 2002) |
| 8Rx9 | *Weeksella zoohelcum* | (Berg et al., 2002) |
| 5Rr4 | *Weeksella zoohelcum* | (Berg et al., 2002) |
| 3Rp5 | *Xenorhabdus luminescens* | (Berg et al., 2002) |
| 3Rp5 | *Xenorhabdus luminescens* | (Berg et al., 2002) |
| 4Rx4 | *Xenorhabdus nematophila* | (Berg et al., 2002) |
| 10Bp14 | *Pantoea agglomerans* | (Berg et al., 2002) |
| 3Bz10 | *Serratia grimesii* | (Berg et al., 2002) |

## 2.3 Bacteriophages

A list of all bacteriophages used in this study is shown in **Table 2.5**. All chloroform used in the sterilisation of phage lysates was saturated with sodium hydrogen carbonate.

### 2.3.1 Isolation and purification of novel environmental phages

Several phages used in this study were newly isolated from the environment. All phages bar one were isolated from water samples from either the Milton sewage works or the river Cam. ΦRC62 was isolated from a rotting potato bought from a supermarket. Briefly the water samples were filtered using a 0.22 µm filter or for ΦRC62 the rot scraped from the potato and suspended in sterile $dH_20$ and filtered. 5 ml of the filtered water samples were mixed in a flask containing 5 ml 2x LB and 500 µl of the host bacterium and enriched overnight at the growth temperature of the host. The following day, 1 ml of this solution was added to a 1.5 ml Eppendorf tube containing 500 µl of chloroform and vortexed for two minutes. It was then centrifuged at 10,000 RPM for 2 minutes. Next, 10 µl of the resulting supernatant was added to 200 µl of an overnight culture of the appropriate host and mixed with 4 ml of molten top agar and poured as a top lawn onto LBA plates. The plates were left to set and incubated overnight at the growth temperature of the host bacteria. A single plaque from plates showing plaques, was picked with a sterile cocktail stick and added to a 1.5 ml Eppendorf tube containing 200 µl of phage buffer. 20 µl of chloroform was added to this and then vortexed quickly. This was spun down and the supernatant serially diluted and the phage plated out to single plaques. Another single plaque was picked and the process repeated three times to ensure phage purity and this phage was then made used to make phage lysates of high titre.

**Table 2.5** Phages used in this study

| Phage | Host isolated on | Relevant characteristics | Reference or source |
|---|---|---|---|
| ΦA1 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦA2 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦA2 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦA4S | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦA4L | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦA5 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦA6 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦM2 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦM4 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦM5 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS21 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS22 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS32 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS33 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS34 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS41 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS42 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS61 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS64 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦS71 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |

Table continued overleaf

**Table 2.5** continued. Phages used in this study

| Phage | Host isolated on | Relevant characteristics | Reference or source |
|---|---|---|---|
| ΦS72 | SCRI1043 | Wild type | Finham sewage works, Conventry (Toth, 1991) |
| ΦB1 | SCRI1043 | Wild type | Gift from Arild Sletten (Bioforsk, Aas, Norway) (Blower et al., 2009) |
| ΦB24 | SCRI1043 | Wild type | Gift from Arild Sletten (Bioforsk, Aas, Norway) (Blower et al., 2009) |
| ΦR1 | SCC34 | Wild type | River Cam, Cambridge (This study) |
| ΦRC6 | SCRI1043 | Wild type | Willow pond, botanical garden, Cambridge (This study) |
| ΦRC9 | SCRI1043 | Wild type | Milton sewage works, Cambridge (This study) |
| ΦRC18 | SCRI1043 | Wild type | Milton sewage works, Cambridge (This study) |
| ΦRC30 | SCRI1043 | Wild type | River Cam, Cambridge (This study) |
| ΦRC33 | SCRI1043 | Wild type | River Cam, Cambridge (This study) |
| ΦRC36 | SCRI1043 | Wild type | River Cam, Cambridge (This study) |
| ΦRC41 | SCRI1043 | Wild type | Milton sewage works, Cambridge (This study) |
| ΦRC50 | SCRI1043 | Wild type | River Cam, Cambridge (This study) |
| ΦRC58 | SCRI1043 | Wild type | River Cam, Cambridge (This study) |
| ΦRC62 | SCRI1043 | Wild type | Rotting potato, “Gift” from George, Cambridge |
| ΦRC63 | SCRI1043 | Wild type | River Cam, Cambridge (This study) |
| ΦRC64 | SCRI1043 | Wild type | River Cam, Cambridge (This study) |
| ΦR11 | SCC34 | Wild type | River Cam, Cambridge (This study) |
| ΦTE | SCC34 | Wild type | (Blower et al., 2012a) |
| ΦM1 | SCRI1043 | Wild type | (Toth et al., 1997) |
| ΦM1-E1 | SCRI1043 (pTA46) | $ToxIN_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E2 | SCRI1043 (pTA46) | $ToxIN_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E3 | SCRI1043 (pTA46) | $ToxIN_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E4 | SCRI1043 (pTA46) | $ToxIN_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E5 | SCRI1043 (pTA46) | $ToxIN_{Pa}$ escape | (Blower et al., 2017) |

Table continued overleaf

**Table 2.5** continued. Phages used in this study

| Phage | Host isolated on | Relevant characteristics | Reference or source |
|---|---|---|---|
| ΦM1-E6 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E7 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E8 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E9 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E10 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E11 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E12 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E13 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E14 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E15 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E16 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E17 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E18 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E19 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E20 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E21 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E22 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E23 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E24 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E25 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E26 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E27 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E28 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E29 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E30 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E31 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E32 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E33 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E34 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E35 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E36 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E37 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E38 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E39 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E40 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E41 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E42 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E43 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E44 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E45 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E46 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |

Table continued overleaf

**Table 2.5** continued. Phages used in this study

| Phage | Host isolated on | Relevant characteristics | Reference or source |
|---|---|---|---|
| ΦM1-E47 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E48 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦM1-E49 | SCRI1043 (pTA46) | ToxIN$_{Pa}$ escape | (Blower et al., 2017) |
| ΦRC10 | SCRI1043 | Wild type | River Cam, Cambridge (This study) |
| ΦRC10a | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10b | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10c | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10d | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10e | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10f | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10g | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10h | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10i | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10j | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10k | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10l | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10m | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10n | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10o | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦRC10p | SCRI1043 (pFR2) | TenpIN$_{Pl}$ escape | This study |
| ΦCHAI1 | DH5α | Wild type | River Cam, Cambridge (This study) |
| ΦCHAI2 | DH5α | Wild type | River Cam, Cambridge (This study) |
| ΦCHAI3 | DH5α | Wild type | River Cam, Cambridge (This study) |
| ΦCHAI4 | DH5α | Wild type | River Cam, Cambridge (This study) |
| ΦCHAI5 | DH5α | Wild type | River Cam, Cambridge (This study) |
| ΦCHAI6 | DH5α | Wild type | River Cam, Cambridge (This study) |
| ΦCHAI7 | DH5α | Wild type | River Cam, Cambridge (This study) |
| ΦCHAI8 | DH5α | Wild type | River Cam, Cambridge (This study) |
| ΦCHAI8a | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦCHAI8b | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦCHAI8c | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦCHAI8d | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦCHAI8e | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦCHAI9 | DH5α | Wild type | River Cam, Cambridge (This study) |

Table continued overleaf

**Table 2.5** continued. Phages used in this study

| Phage | Host isolated on | Relevant characteristics | Reference or source |
|---|---|---|---|
| ΦCHAI9-E1 | EPI300 (pFR2) | TenpIN$_{PI}$ escape | This study |
| ΦTRB18 | DH5α | Wild type | (Blower, 2009) |
| ΦTRB19 | DH5α | Wild type | (Blower, 2009) |
| ΦF1 | DH5α | Wild type | (Rao, 2014) |
| ΦF2 | DH5α | Wild type | (Rao, 2014) |
| ΦF4 | DH5α | Wild type | (Rao, 2014) |
| ΦF4-E1 | EPI300 (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦF4-E2 | EPI300 (pFR2) | TenpIN$_{PI}$ escape | This study |
| ΦF5 | DH5α | Wild type | (Rao, 2014) |
| ΦF6 | DH5α | Wild type | (Rao, 2014) |
| ΦF7 | DH5α | Wild type | (Rao, 2014) |
| T4 | *E. coli* | Wild type | Laboratory collection |
| T6 | *E. coli* | Wild type | Laboratory collection |
| T6-E1 | EPI300 (pFR2) | TenpIN$_{PI}$ escape | This study |
| T6-E2 | EPI300 (pFR2) | TenpIN$_{PI}$ escape | This study |
| T6-E3 | EPI300 (pFR2) | TenpIN$_{PI}$ escape | This study |
| T6-E4 | EPI300 (pFR2) | TenpIN$_{PI}$ escape | This study |
| T6-E5 | EPI300 (pFR2) | TenpIN$_{PI}$ escape | This study |
| T6-E6 | EPI300 (pFR2) | TenpIN$_{PI}$ escape | This study |
| ΦOT8s | *S* 39006 (pMUT13) | Wild type | (Blower, 2009) |
| ΦOT8s-E1 | *S* 39006 (pMUT13, pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E2 | *S* 39006 (pMUT13, pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E3 | *S* 39006 (pMUT13, pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E4 | *S* 39006 (pMUT13, pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E5 | *S* 39006 (pMUT13, pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E6 | *S* 39006 (pMUT13, pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E7 | *S* 39006 (pMUT13, pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E8 | *S* 39006 (pMUT13, pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E9 | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E10 | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E11 | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E12 | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E13 | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |
| ΦOT8s-E14 | DH5α (pTA46) | ToxIN$_{Pa}$ escape | This study |

Table continued overleaf

**Table 2.5** continued. Phages used in this study

| Phage | Host isolated on | Relevant characteristics | Reference or source |
|---|---|---|---|
| ΦOT8s-E15 | DH5α (pTA46) | $ToxIN_{Pa}$ escape | This study |
| ΦOT8s-E16 | DH5α (pTA46) | $ToxIN_{Pa}$ escape | This study |
| ΦLT | TT01 | Wild type | Feng Rao (unpublished) |
| ΦJB2 | TT01 | Wild type | Jessica Bergman (unpublished) |
| ΦJB15 | TT01 | Wild type | Jessica Bergman (unpublished) |
| ΦRAY0 | TT01 | Wild type | This study |
| ΦRAY1 | TT01 | Wild type | This study |
| ΦRAY19 | TT01 | Wild type | This study |
| ΦRAY26 | TT01 | Wild type | This study |
| ΦRAY27 | TT01 | Wild type | This study |
| ΦYSD1 | 5383 | Wild type | This study |

### 2.3.2 Isolation of TA escape phages

The wild type Abi-sensitive phages were plated out on the appropriate lawn of bacteria containing the Type III TA plasmid of choice. Briefly, this involved 10 µl of serially diluted plaque pure phage onto 100 µl of overnight bacteria mixed with 4 ml molten top agar, poured as an overlay onto LBA plates. Plates were left to cool and incubated at the growth temperature of the host. A single plaque was picked from the resulting plaques using a sterile cocktail stick and added to a 1.5 ml Eppendorf tube containing 200 µl of phage buffer. 20 µl of chloroform was then added, the sample vortexed and spun down. 10 µl of the supernatant was then used to make a serial dilution and a phage lysate was made as previously described using the same Type III TA containing host as before. Escapes were confirmed by efficiency of plating assays using a functioning and non-functioning frameshift Type III TA system containing strains.

### 2.3.3 Determining phage titres and production of phage lysates

Phage titres were calculated by plaque forming units per ml of phage lysate. To determine the titre, the lysate was serially diluted in phage buffer. Next 10 µl of each dilution was added to 200 µl of the appropriate overnight culture of bacteria and mixed with 4 ml of molten top agar. This mixture was poured evenly onto LBA plates and left to set. Plates were then incubated overnight at the growth temperature of the host bacteria. The next day plaques were counted to determine the titre.

To make phage lysates, plates with plaques that were semi-confluent i.e. overlapping plaques with small areas of bacterial lawn still remaining were used to generate the lysates. The top lawn was scraped off into a glass universal containing 2 ml of phage buffer and 500 µl of chloroform. This was then vortexed for 2 mins and then centrifuged at 4500 RPM for 20 mins at 4ºC. The resulting supernatant was then pipetted into a glass bijou bottle with 100 µl of chloroform added to maintain sterility. These phage lysates were stored at 4ºC.

### 2.3.4 Determining efficiency of plating

The efficiency of plating (EOP) is used as a measure of how well a phage replicates in one host compared to another (Ellis and Delbrück, 1939). EOPs were used for determining the strength of abortive infection by the Type III TA systems, as well as in

host range testing. For the Type III TA testing, the phage titre on a host containing a functional Type III TA system was divided by the phage titre on the same host containing a non-functioning frameshift version. With regards to the host range testing, the EOP is the titre of the phage on the new bacterial host divided by the titre of the phage on the host that the phage lysate was made with. Thus an EOP of 1 means that there is no difference in replication on the test host compared to the control host. EOPs greater than 1 mean that replication is more productive in the test host and EOPs less than 1 signifies weaker replication in the test host.

### 2.3.5 Adsorption assays

Adsorption assays were performed using the *Photorhabdus luminescens* TT01 phage, ΦRAY26. For these assays, an overnight of bacterial culture was added to a 25 ml flask up to a volume of 10 ml and adjusted to an $O.D._{600}$ reading of 1. The phage, ΦRAY26 was added at a multiplicity of infection (MOI) of 0.01. The MOI is calculated as the titre of phage in $pfu.ml^{-1}$ divided by number of bacterial cells as cells per ml. The flasks were added to a water bath with shaking at 300 RPM and at 30ºC. As soon as phage was added a 100 µl sample was taken from the flask as a time zero sample. This was added to a 1.5 ml Eppendorf tube containing 900 µl of phage buffer and 50 µl chloroform, and vortexed briefly then put on ice for the duration of the assay. The assay was run for 30 mins with samples taken at 2.5 min intervals for the first 15 mins, then a sample at 20 mins and a final sample at 30 mins. All samples taken were immediately vortexed in the same conditions as the time zero sample. After the assay, samples were centrifuged for 5 mins at 12,000 RPM. 10 µl of the resulting supernatant was used to determine the titre of the phage. Results were recorded as the ratio of free phage for each time point compared to time zero.

### 2.3.6 Transmission electron microscopy of phages

Phages to be imaged by transmission electron microscopy were first grown to a high titre lysate using top agarose. 10 µl of this phage lysate was added to a charged copper grid for 5 mins. The grids were then washed twice in filtered $dH_20$ by placing the grids on a drop of water and drying on filter paper immediately. Grids were then negatively stained for 2 mins using 2% uranyl acetate. The electron microscope used for taking phages images was a Tecnai G2 series transmission electron microscope.

## 2.4 Recombinant DNA techniques

### 2.4.1 Bacterial plasmid extraction

Extraction of plasmids were performed using the Thermo scientific miniprep kit. Protocol used as stated by the manufacturers. After extraction, plasmids were either used immediately or stored at -20ºC. A list of all plasmids used is shown in **Table 2.6**.

**Table 2.6** Plasmids used in this study

| Plasmid | Description | Reference or source |
|---|---|---|
| pTA46 | *toxI*$_{Pa}$, *toxN*$_{Pa}$ with native promoter in pBR322, Ap$^R$ | (Fineran et al., 2009) |
| pTA47 | *toxI*$_{Pa}$, *toxN*$_{Pa}$-FS with native promoter in pBR322, Ap$^R$ | (Fineran et al., 2009) |
| pFR2 | *tenpI*$_{Pl}$, *tenpN*$_{Pl}$ with native promoter in pBR322, Ap$^R$ | (Blower et al., 2012b) |
| pFR8 | *tenpI*$_{Pl}$, *tenpN*$_{Pl}$-FS with native promoter in pBR322, Ap$^R$ | (Blower et al., 2012b) |
| pTA114 | *toxI*$_{Bt}$, *toxN*$_{Bt}$ with native promoter in pBR322, Ap$^R$ | (Short et al., 2013) |
| pFLS63 | *toxI*$_{Bt}$, *toxN*$_{Bt}$-FS with native promoter in pBR322, Ap$^R$ | (Short et al., 2013) |
| pFLS116 | *cptI*$_{Er}$, *cptN*$_{Er}$ with native promoter in pBR322, Ap$^R$ | (Blower et al., 2012b) |
| pTW03 | *cptI*$_{Er}$, *cptN*$_{Er}$-FS with native promoter in pBR322, Ap$^R$ | Tom Whittaker, unpublished |
| pMUT13 | *lamB*, Km$^R$ | (McGowan et al., 2005) |
| pTRC99a | Ap$^R$ | Gift from, David Clarke's lab |
| pBMM800 | TT01 *flgG* in pTRC99a, Ap$^R$ | (Easom and Clarke, 2008) |
| pBMM802 | TT01 *motAB* in pTRC99a, Ap$^R$ | (Easom and Clarke, 2008) |

### 2.4.2 Phage gDNA extraction

Phage genomic DNA was extracted by a phenol/chloroform DNA extraction method with the use of vacuum grease to separate the organic and aqueous phases. Phages selected for genomic extraction were made as high titre lysates using top agarose to prevent impurities from agar. For each extraction, three 1.5 ml Eppendorf tubes were filled with a pea size amount of sterile vacuum grease (Dow Corning). These tubes were centrifuged at 4500 RPM for 1 min to pellet the grease. In one of these tubes, 450 µl of phage lysate, 4.5 µl of 1mg/ml DNase I and 2.5 µl of 10 mg/ml RNase were added. This tube was then incubated for 30 mins at 37ºC. Next 11.5 µl 20% SDS and 4.5 µl of 10mg/ml proteinase K were added to the tube and incubated for a further 30

mins at 37ºC. After incubation, 500 µl of a phenol:chloroform:isoamyl alcohol (25:24:1) mixture was added. The tube was inverted three times and centrifuged for 5 mins at 4500 RPM. The resulting tube should have a top aqueous phase that is separated from the organic phase by a thin layer of vacuum grease. The supernatant (aqueous phase) was removed and added to a new 1.5 ml Eppendorf tube containing vacuum grease. 500 µl of the phenol:chloroform:isoamyl mixture was added, mixed and centrifuged as before. The resulting top supernatant was transferred again to the final 1.5 ml Eppendorf tube containing vacuum grease and 500 µl of a chloroform:isoamyl alcohol (24:1) mixture was added. This was inverted three times to mix and then centrifuged for 5 mins at 4500 RPM. The supernatant was then transferred to a new 1.5 ml Eppendorf tube with no vacuum grease. 4.5 µl of 3M sodium acetate adjusted to pH 5.2 and 500 µl 100% isopropanol was added. The tube was then incubated for at least 15 mins at room temperature. Next the tubes were centrifuged at 12000 RPM for 20 mins at 4ºC to pellet the precipitated DNA. The supernatant was discarded and the pellet was washed twice with 400 µl 70% ethanol by centrifugation at 12000 RPM for 5 mins and discarding the supernatant. After the final wash the pellet was dried on the bench to remove residual ethanol and then suspended in 100 µl $dH_20$. DNA was stored at 4ºC for short term use or frozen at -20ºC for long term storage.

### 2.4.3 Gel electrophoresis

Gel electrophoresis was used to check the quality of genomic DNA, plasmids and PCR amplified products. The agarose gels were made by adding agarose in TAE buffer and dissolved using a microwave. After ensuring the agarose was completely dissolved, the solution was left to cool for 5 mins after which ethidium bromide was added. Gels were run in TAE buffer at 80 volts for 40 mins. For greater resolution a lower voltage and longer time was used. Gels were visualised using the Gel view UV transilluminator connected to the Gene Genious Bio-Imaging system from Syngene synoptics.

### 2.4.4 DNA extraction from gels

DNA bands from gels that required further manipulation were cut from gels using a scalpel. The DNA was then extracted from the band using the GeneJET gel extraction kit (ThermoFisher Scientific) according to the protocol from the manufacturer.

### 2.4.5 Picogreen assay

To determine the purity and concentration of dsDNA from phage genomic DNA extractions, a picogreen assay was used. The kit used was the Quanti-iT™ PicoGreen ™ dsDNA assay kit (ThermoFisher Scientific). The protocol used was according to the manufacturer's instructions.

### 2.4.6 Polymerase chain reaction

All PCR reactions were performed using the Phusion polymerase (NEB). The components required for the PCR reaction are shown in **Table 2.7**. PCR reactions were run using the Veriti 96 well thermal cycler (Applied Biosystems) with the conditions shown in **Table 2.8**. All primers used in this study are shown in **Table 2.9**.

**Table 2.7** PCR components

| Component | Volume (µl) |
|---|---|
| Forward primer | 2.5 |
| Reverse primer | 2.5 |
| 5x HF buffer | 10 |
| DNA | 1 |
| Phusion polymerase | 0.5 |
| $dH_20$ | Up to 50 µl |

**Table 2.8** PCR conditions

| Condition | Temperature (ºC) | Time |
|---|---|---|
| Initial denaturation | 98 | 2 mins |
| Cycle 29 times | | |
| Denaturation | 98 | 2 sec |
| Annealing | 55-60 (varies depending on primers) | 15 sec |
| Extension | 72 | 15 sec per kb |
| Final extension | 72 | 2 mins |
| Hold | 4 | Indefinite |

**Table 2.9** Primers used in this study

| Primer | Sequence (5'-3') | Description |
|---|---|---|
| TRB35 | AAAGTGCCACCTGACGTC | Forward sequencing primer for pBR322 |
| TRB36 | ATGACGATGAGCGCATTG | Reverse sequencing primer for pBR322 |
| oRC19 | ACCTACAATGCCCCAGATGC | Forward sequencing primer for *phim1-23* and *rc10-26* |
| oRC20 | CGGTCGTACTTGGCTTCGGG | Reverse sequencing primer for *phim1-23* and *rc10-26* |
| oRC34 | GAAGTCTACTGATGTTGTGA | Forward sequencing primer for ΦOT8s *motA* |
| oRC35 | GAATAACCGGTTGTGACTGC | Reverse sequencing primer for ΦOT8s *motA* |
| oRC52 | GAGTCATAATCTTCAGAACG | Forward sequencing primer for ΦOT8s *asiA* |
| oRC53 | AACTGTGACGATGTATTGGT | Reverse sequencing primer for ΦOT8s *asiA* |
| oRC54 | GTACGCTGACGATGAATATT | Forward sequencing primer for T6 *alc* |
| oRC55 | TCTTCTGCGAACTTGTTGAA | Reverse sequencing primer for T6 *alc* |

### 2.4.7 DNA sequencing

#### 2.4.7.1 Sequencing of plasmids and PCR amplicons

All sequencing of plasmids and PCR amplified products were performed at the GATC now part of Eurofins Genomics in Germany. 7.5 µl of the DNA sample was sent with 2.5 µl of the primer in a 1.5 ml Eppendorf tube.

#### 2.4.7.2 Whole genome sequencing

Prior to sequencing all extracted phage genomes were run on agarose gels and a picogreen assay performed to confirm concentration and quality. All phages, apart from five TT01 phages, were sent for whole genome sequencing at MicrobesNG based in Birmingham. The platform used for sequencing at MicrobesNG is Illumina using 2x250 bp paired-end reads and contigs were assembled using SPAdes. Due to the small size of phage genomes the majority of genomes were returned as a single contig. The quality of genomes varied but generally coverage was several times that of the genome with certain genome coverages reaching as high as 1000x. The data for all phage genomes including FASTA sequences, annotations and coverage is available in the appendices.

As mentioned above, five TT01 phage genomes were not sequenced by MicrobesNG and were sent for sequencing at the Sanger Institute, located at the genome campus in Hinxton. These genomes were sequenced by Pacbio sequencing twice and genomes were assembled using contigs returned from both sequencing runs. As before, all information for these genomes is available in the appendices.

#### 2.4.7.3 Genome annotations and analysis

Phage genomes were annotated using the Artemis genome browser that is freely available from the Wellcome Sanger Institute website. For the prediction and annotation of open reading frames, two gene prediction tools, Glimmer and GeneMark were used. Protein function and homologous proteins were identified by performing a protein BLAST on all predicted proteins. For the identification of tRNA genes, tRNAScan and ARAGORN was used. Nucleotide comparisons were performed using EMBOSS stretcher with Artemis comparison tool used for protein alignments. For the construction of phylogenetic trees MEGA7 was used and proteins were aligned using muscle.

### 2.4.8 Bacterial transformation

#### 2.4.8.1 Chemical transformation

Chemical or heat shock transformation of plasmid DNA was used for *E. coli* cells. To prepare the competent cells, 500 µl of the bacterial overnight was grown in 25 ml LB supplemented with 375 µl of 1 M $MgCl_2$. The culture was grown in a 250 ml flask with shaking at 250 RPM at 37ºC until the $O.D._{600}$ was between 0.4 to 0.6. The culture was then added to a pre-chilled universal tube and left for 1 hour at 4ºC. This was then centrifuged at 4500 RPM for 10 mins at 4ºC. The supernatant was discarded and the resulting pellet suspended in 10 ml of solution A that had been chilled on ice. This was then left on ice for 20 mins then centrifuged under the same conditions as before. The supernatant was discarded again and the pellet suspended in 1 ml of solution A 15% glycerol to make the competent cell mixture. This mixture was then used immediately or stored at -80ºC for later use.

For transformation 50 µl of the competent cell mixture was added to a 1.5 ml Eppendorf tube with 0.5 µl of miniprepped plasmid DNA and left on ice for 1 hour. This was then moved to a water bath at 42ºC for 2 mins after which 1 ml of LB was added and the cells were left to recover on a rotating wheel for 1 hour at 37ºC. The cells were then spread onto LBA plates containing the appropriate selection and incubated at 37ºC overnight.

#### 2.4.8.2 Electroporation transformation

For non *E. coli* cells, electroporation was used. The standard electroporation protocol is shown below, for *Ph. luminescens* an adapted protocol was used. To prepare electroporation competent cells, a 250 ml flask containing 25 ml of LB was inoculated with 500 µl of an overnight culture of the bacteria. This was grown in a water bath at the growth temperature of the bacteria, with shaking at 250 RPM. The bacterial culture was grown to an $O.D._{600}$ of 0.6 then decanted into a pre-chilled universal tube and left on ice for 30 mins. The cells were then pelleted by centrifugation at 4500 RPM for 10 mins at 4ºC. The resulting supernatant was discarded and the pellet suspended in 10 ml ice cold $dH_20$. It was immediately pelleted again as before and suspended in 10 ml ice cold 10% glycerol. This was also pelleted and the pellet finally suspended in 1 ml 10% glycerol. These electrocompetent cells were either used immediately or stored at -80ºC for later use.

For the electroporation, 50 µl of electrocompetent cells and 1 µl of plasmid was added to a precooled electroporation cuvette. Conditions for electroporation were 200Ω, 25 µFd and 2.5 kV. Cells were left to recover for 1 hr in 1 ml of LB at 30ºC and then spread on LBA plates containing the appropriate antibiotic. The DNA from resulting colonies was sequenced to confirm the successful uptake of the plasmid and to confirm that no mutations had occurred.

#### 2.4.8.3 *Ph. luminescens* electroporation

2 ml of an overnight culture of *Ph. luminescens* were added to 200 ml LB containing 0.1% sodium pyruvate. The culture was grown in a 2 litre flask with shaking at 250 RPM at 30ºC until reaching an $O.D._{600}$ reading of 0.3. It was then chilled on ice for 90 mins after which the culture was pelleted by centrifugation at 4500 RPM for 20 mins at 4ºC. The supernatant was discarded and the pellet suspended in 200ml of *Photorhabdus* wash buffer. This wash was repeated twice but suspended in 100 ml then 10 ml wash buffer. Finally, after the last spin, the cells were suspended in 200 µl wash buffer and were used immediately. In a prechilled electroporation cuvette, 50 µl of competent *Ph. luminescens* and 1.5 µl of plasmid DNA were added. The cells were electroporated at 2.5 kV, 200 Ω and 25 µF with a time constant of 5 ms. Cells were recovered immediately by the addition of 1 ml LB and transferred to a 1.5 ml Eppendorf tube to recover for 3 hours at 30ºC. Selection of transformant cells was performed by plating the recovered cells onto LBA plates containing 0.1% sodium pyruvate and the appropriate antibiotic. Plates were incubated for 3 days at 30ºC and the DNA from any resulting colonies was sequenced to confirm the uptake of the plasmid.

# Chapter 3 Isolation and characterisation of *Pectobacterium atrosepticum* phages

## 3.1 Introduction

To date there are only five phages isolated on *P. atrosepticum*, including ΦTE and ΦM1 that have had their genomes entirely sequenced. Based on sequence homology they can be split into two groups, either ΦTE-like or ΦM1-like. The ΦTE-like group contains two members, ΦTE and vB_PatM_CB7 (direct submission, CB7 for short, accession no. KY514263). The ΦM1-like group contains three members, ΦM1, Peat1 (Kalischuk et al., 2015)(accession no. KR604693) and PP90 (direct submission, accession no. NC_031096). *P. atrosepticum* causes black leg in potatoes and as such the use of phages have been studied as a potential solution to prevent the disease (Buttimer et al., 2017a). In fact, it has been reported that Tesco has been using the bacteriophage cocktail, Biolyse, developed by APS biocontrol to reduce soft rot and enhance the shelf life of their potatoes, prior to packaging (Branston, 2012). This bacteriophage cocktail will undoubtedly include several uncharacterised *P. atrosepticum* phages and as they will still be viable on potatoes bought in supermarkets, it would be useful to characterise these phages further. At least two of the five fully sequenced *P. atrosepticum* phages, ΦTE and ΦM1 have been shown to be transducers. While this makes these phages useful molecular tools, it makes them undesirable to be used as part of a phage cocktail since they may cause the unwanted horizontal gene transfer of genes involved in antibiotic resistance and virulence. The limited number of *P. atrosepticum* phage genomes published, the current use of these phages in commericial phage cocktails and the fact that these phages could be transducers necessitates further characterisation. Furthermore, genomic data on new *P. atrosepticum* phages isolated in this laboratory will assist with defining the natures of the phages that are aborted and or not aborted by the Type III TA systems.

## 3.2 Whole genome sequencing of 36 novel environmental *P. atrosepticum* phages

To address the paucity of information that there is on *P. atrosepticum* phages, 36 phages isolated on *P. atrosepticum* SCRI1043 or SCC34 were sent for whole genome

sequencing. These phages were isolated over a period ranging from June 1987 to October 2017, derived from water samples obtained from Coventry, Cambridge or Norway. One phage, ΦRC62, was isolated during this period from a rotting potato originating from a supermarket in Cambridge. 14 of these phages sequenced were from a library of phages isolated during this study. These are phages with the prefix, ΦRC or ΦR followed by a number. The remaining phages were isolated by past members of the lab or collaborators. Genomic data returned for these 36 phages showed that they can be split into six groups based on sequence homologies. The groups are as follows, ΦTE-like, ΦM1-like (based on the previously published ΦTE and ΦM1 phages) and four new groups named after the first phage isolated in each group called ΦR1-like, ΦRC6-like, ΦRC63-like and ΦR11-like.

### 3.2.1 ΦTE-like phages

#### 3.2.1.1 Genomic analysis of ΦTE-like phages

Of the 36 *P. atrosepticum* phages sequenced, 12 showed high similarity to ΦTE. As such, these phages belong to the genus *Cr3virus* derived from the *Vequintavirinae* subfamily. These 12 phages can be further spilt into three sub groups based on percentage identity of aligned sequences (**Table 3.1**). Within each subgroup, phages are highly similar to each other, sharing more than 99% identity. Subgroup one comprises two members, ΦS34 and ΦS61, and shares the highest identity with ΦTE. Subgroup two which consists of three members, (ΦA2, ΦA5 and ΦS22) also shares a high sequence identity with ΦTE and with members of subgroup one, but to a lesser degree. Subgroup three is the largest group, consisting of the remaining seven phages and is the least similar to ΦTE of the subgroups. Phage CB7 was also added to the analysis and is the least related to all other ΦTE-like phages. Representatives from each subfamily, ΦS34, ΦA2 and ΦA4S were annotated. Analysis of these phages showed that ΦTE-like phages have genome lengths of 142,349 to 145,315 base pairs with 242 to 266 predicted open reading frames (ORFs) (**Table 3.2**). The ΦTE-like phages also carry at least 2 tRNAs, apart from CB7 which only has a single tRNA. All the phages carry tRNA-Tyr while the second tRNA, if present, is either tRNA-Cys or tRNA Leu. A genomic map of ΦS34 is shown in **Figure 3.1**. ACT alignments made using a tblastx comparison show that very small differences exist within subgroups but

outside of subgroups there are less conserved proteins spread throughout the genomes (**Figure 3.2**).

**Table 3.1** Sequence identity between ΦTE-like phages (%)

| | | Subgroup 1 | | Subgroup 2 | | | Subgroup 3 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Phage | TE | S34 | S61 | A2 | A5 | S22 | A4S | A4L | A6 | S21 | S41 | S72 | B1 | CB7 |
| TE | 100 | 94.9 | 94.9 | 94.6 | 94.6 | 94.6 | 88.9 | 88.3 | 88.9 | 89.9 | 88.3 | 88.9 | 89 | 87.1 |
| S34 | 94.9 | 100 | 99.8 | 93.5 | 93.3 | 93.3 | 87.5 | 86.9 | 87.6 | 87.5 | 86.9 | 87.5 | 87.5 | 84.6 |
| S61 | 94.9 | 99.8 | 100 | 93.5 | 93.3 | 93.3 | 87.5 | 87 | 87.6 | 87.5 | 86.9 | 87.5 | 87.6 | 84.6 |
| A2 | 94.6 | 93.5 | 93.5 | 100 | 99.8 | 99.8 | 89.7 | 89.1 | 89.7 | 89.7 | 89.2 | 89.7 | 89.8 | 85 |
| A5 | 94.6 | 93.3 | 93.3 | 99.8 | 100 | 99.8 | 89.7 | 89.1 | 89.7 | 89.7 | 89.2 | 89.7 | 89.8 | 84.8 |
| S22 | 94.6 | 93.3 | 93.3 | 99.8 | 99.8 | 100 | 89.7 | 89.1 | 89.7 | 89.7 | 89.2 | 89.7 | 89.8 | 84.8 |
| A4S | 88.9 | 87.5 | 87.5 | 89.7 | 89.7 | 89.7 | 100 | 99.3 | 99.9 | 99.8 | 99 | 99.8 | 99.8 | 83.3 |
| A4L | 88.3 | 86.9 | 87 | 89.1 | 89.1 | 89.1 | 99.3 | 100 | 99.2 | 99.1 | 99.8 | 99.1 | 99.1 | 83.7 |
| A6 | 88.9 | 87.6 | 87.6 | 89.7 | 89.7 | 89.7 | 99.9 | 99.2 | 100 | 99.7 | 99 | 99.8 | 99.8 | 83.3 |
| S21 | 89.9 | 87.5 | 87.5 | 89.7 | 89.7 | 89.7 | 99.8 | 99.1 | 99.7 | 100 | 99 | 99.8 | 99.8 | 83.2 |
| S41 | 88.3 | 86.9 | 86.9 | 89.2 | 89.2 | 89.2 | 99 | 99.8 | 99 | 99 | 100 | 99 | 99 | 83.6 |
| S72 | 88.9 | 87.5 | 87.5 | 89.7 | 89.7 | 89.7 | 99.8 | 99.1 | 99.8 | 99.8 | 99 | 100 | 99.8 | 83.2 |
| B1 | 89 | 87.5 | 87.6 | 89.8 | 89.8 | 89.8 | 99.8 | 99.1 | 99.8 | 99.8 | 99 | 99.8 | 100 | 83.3 |
| CB7 | 87.1 | 84.6 | 84.6 | 85 | 84.8 | 84.8 | 83.3 | 83.7 | 83.3 | 83.2 | 83.6 | 83.2 | 83.3 | 100 |

The colours of the table were generated using conditional formatting in Microsoft excel where the higher the sequence identity the greener the colour is. As the sequence identity lowers the colours change to yellow, orange and finally red which signifies the lowest % identity.

**Table 3.2** Genomic characterisation of ΦTE-like phages

| Phage | Subgroup | Genome size (bp) | Predicted ORFs | tRNAs |
|---|---|---|---|---|
| ΦTE | - | 142,349 | 242 | tRNA-Cys, tRNA-Tyr |
| ΦS34 | 1 | 145,315 | 261 | tRNA-Cys, tRNA-Tyr |
| ΦS61 | 1 | 145,315 | NA | NA |
| ΦA2 | 2 | 144,475 | 266 | tRNA-Leu, tRNA-Tyr |
| ΦA5 | 2 | 144,476 | NA | NA |
| ΦS22 | 2 | 144,481 | NA | NA |
| ΦA4S | 3 | 142,686 | 259 | tRNA-Leu, tRNA-Tyr |
| ΦA4L | 3 | 143,656 | NA | NA |
| ΦA6 | 3 | 142,681 | NA | NA |
| ΦS21 | 3 | 142,685 | NA | NA |
| ΦS41 | 3 | 143,587 | NA | NA |
| ΦS72 | 3 | 142,686 | NA | NA |
| ΦB1 | 3 | 142,686 | NA | NA |
| CB7 | - | 142,778 | 253 | tRNA-Tyr |

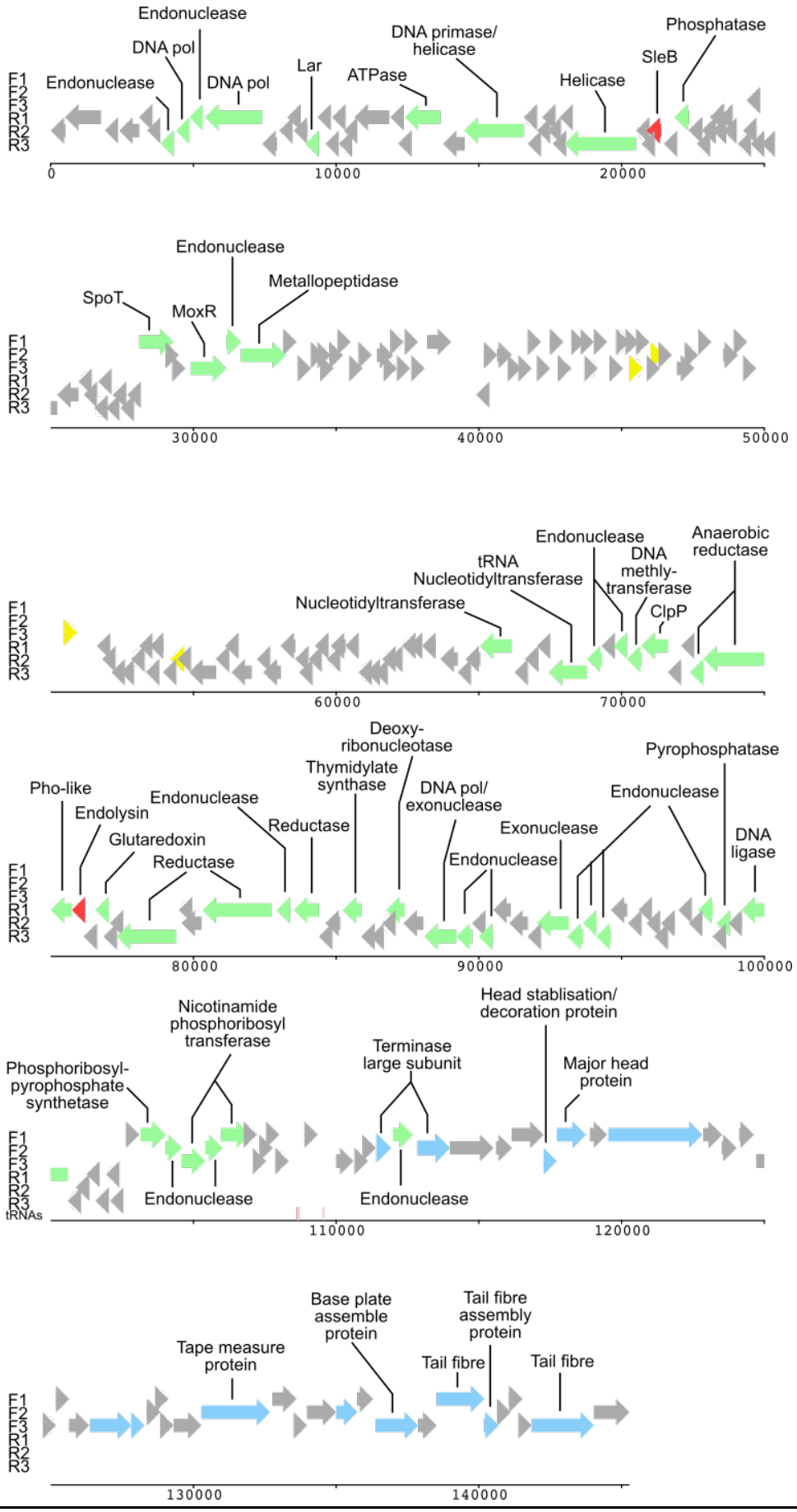

Endonuclease
DNA pol
Endonuclease
DNA pol
Lar
ATPase
DNA primase/
helicase
Helicase
SleB
Phosphatase
F1
F2
F3
R1
R2
R3
0
10000
20000
Endonuclease
SpoT
MoxR
Metallopeptidase
30000
40000
50000
Endonuclease
DNA
methly-
transferase
Anaerobic
reductase
tRNA
Nucleotidyltransferase
Nucleotidyltransferase
ClpP
60000
70000
Deoxy-
ribonucleotase
Thymidylate
synthase
Pyrophosphatase
Pho-like
Endonuclease
DNA pol/
exonuclease
Endonuclease
Endolysin
Reductase
Exonuclease
DNA
ligase
Glutaredoxin
Reductase
Endonuclease
80000
90000
100000
Nicotinamide
phosphoribosyl
transferase
Head stablisation/
decoration protein
Terminase
large subunit
Major head
protein
Phosphoribosyl-
pyrophosphate
synthetase
Endonuclease
Endonuclease
tRNAs
110000
120000
Base plate
assemble
protein
Tail fibre
assembly
protein
Tape measure
protein
Tail fibre
Tail fibre
130000
140000

**Figure legend on next page**

**Figure 3.1** Genomic map of ΦTE-like phage, ΦS34.

Forward reading frames are labelled F1-3 with reverse reading frames labelled as R1-3. Genes involved in metabolism, virion structure or lysis are in green, blue or red, respectively. Genes that have no predicted function but are homologous to hypothetical proteins in previously published phage genomes are shown in grey. Hypothetical protein that have no significant matches to published proteins are shown in yellow. The location of the two tRNAs detected are shown as pink bars.

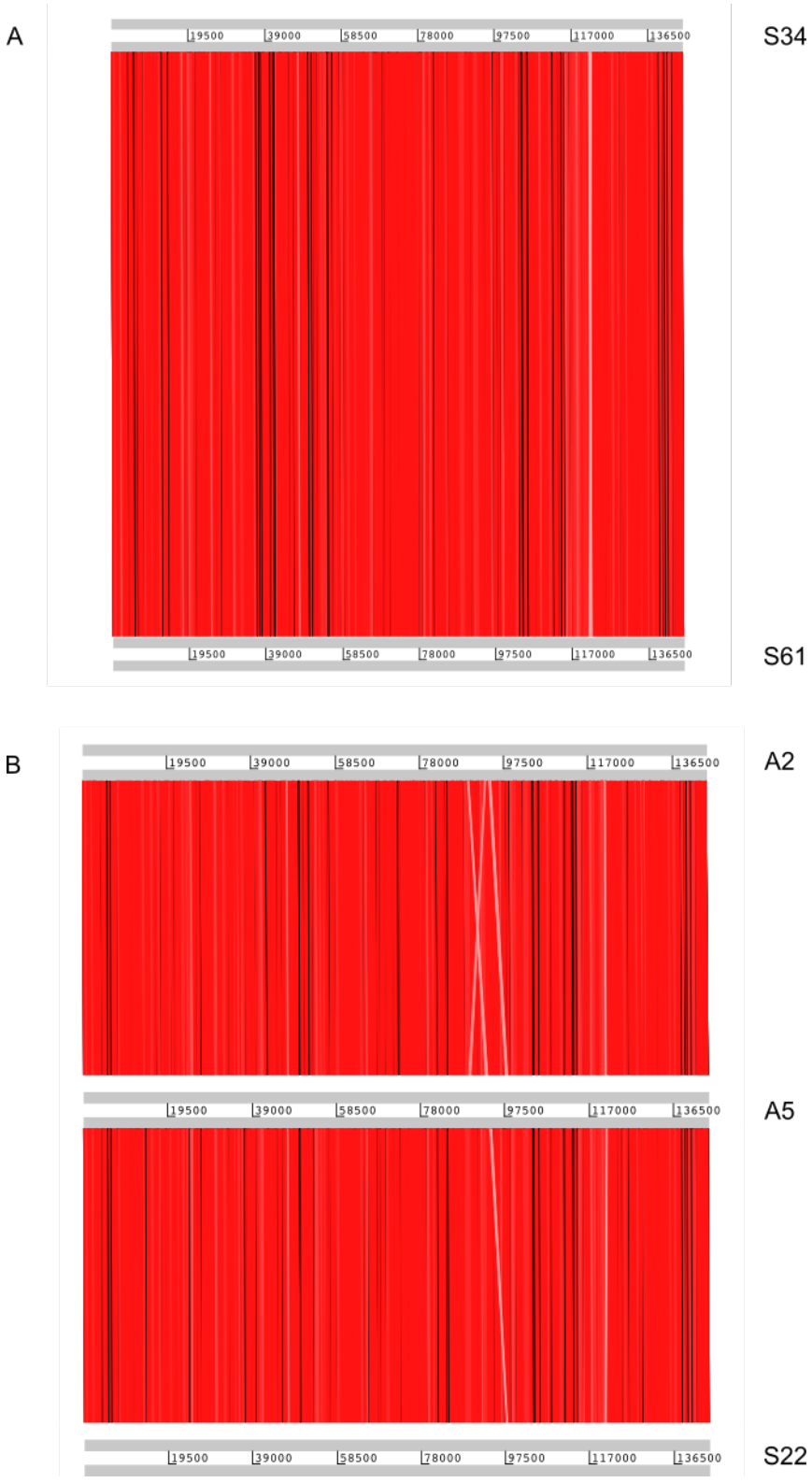



**Figure 3.2** continued overleaf

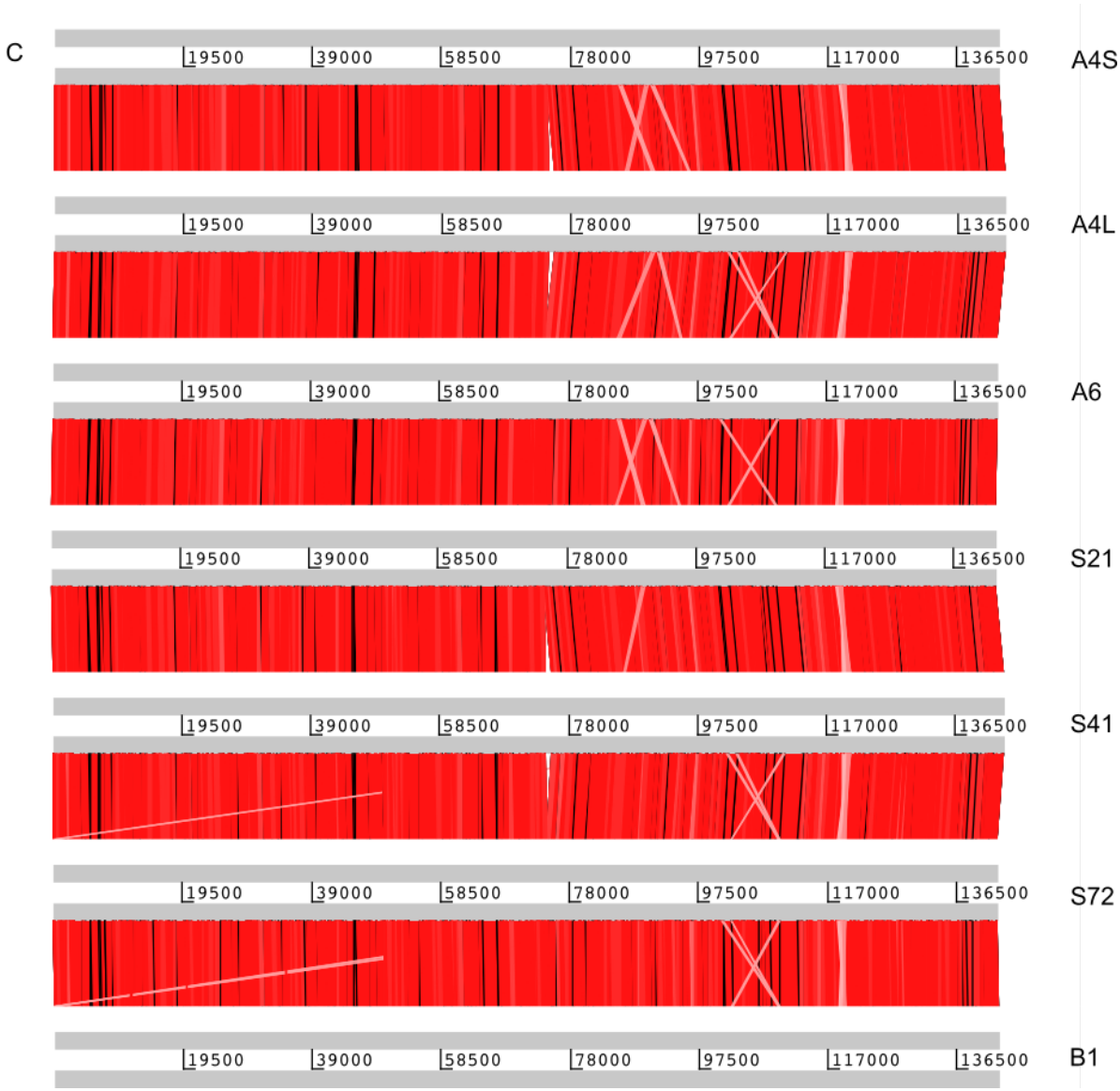



**Figure 3.2** continued overleaf

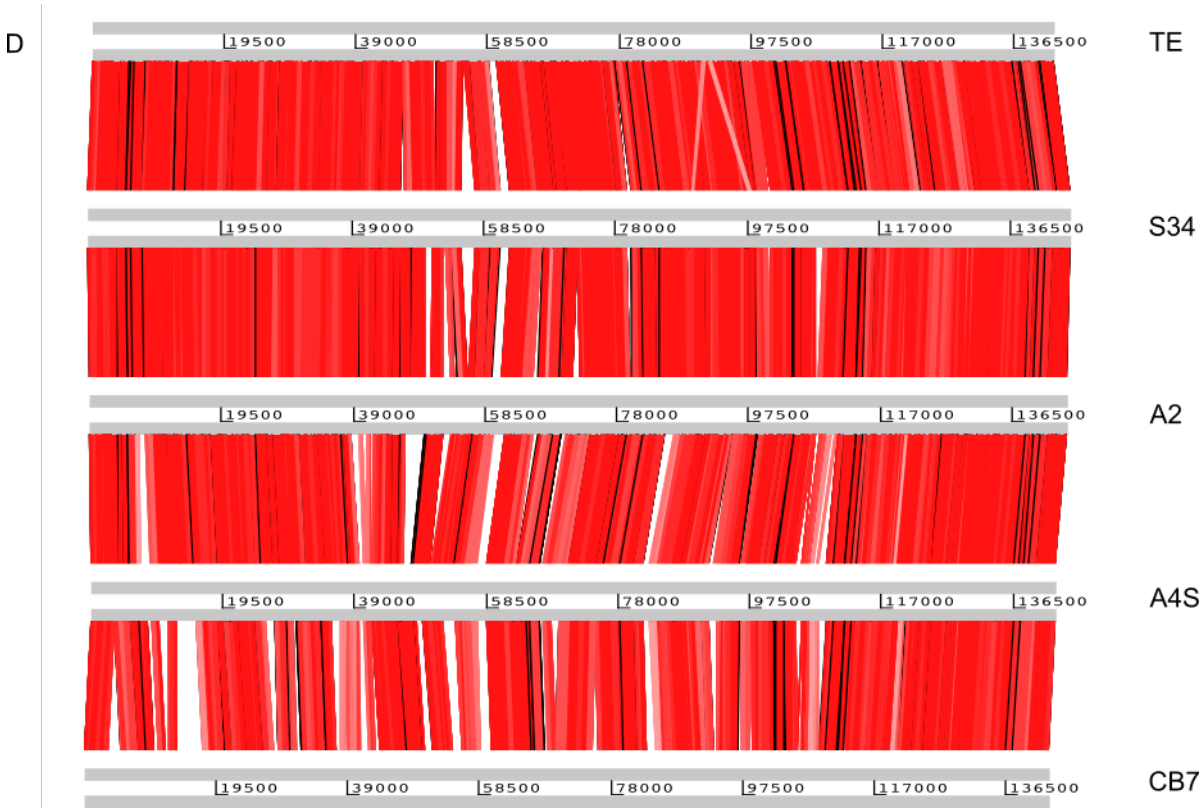



**Figure 3.2** ACT comparison diagrams of ΦTE-like phages.

Genome comparisons of the subgroups one (**A**), two (**B**) and three (**C**) show little differences compared to each other. However outside of the subgroups (**D**) there are more pronounced differences. The comparisons were done using a tblastx comparison file. Red bars show areas of high amino acid similarity.

#### 3.2.1.2 TEM imaging of ΦTE-like phages

All ΦTE-like phages were imaged by transmission electron microscopy. They were all shown to belong to the *Myoviridae* family as characterised by long contractile tails (**Figure 3.3, 3.4**). This is in agreement with the genomic data obtained and images of the phages closely resemble previous TEM images of ΦTE. All the ΦTE-like phages have isometric icosahedral heads measuring roughly 100 nm across from either side of the icosahedral face with extended tails of about 120 nm. No contracted phages were found during the imaging.

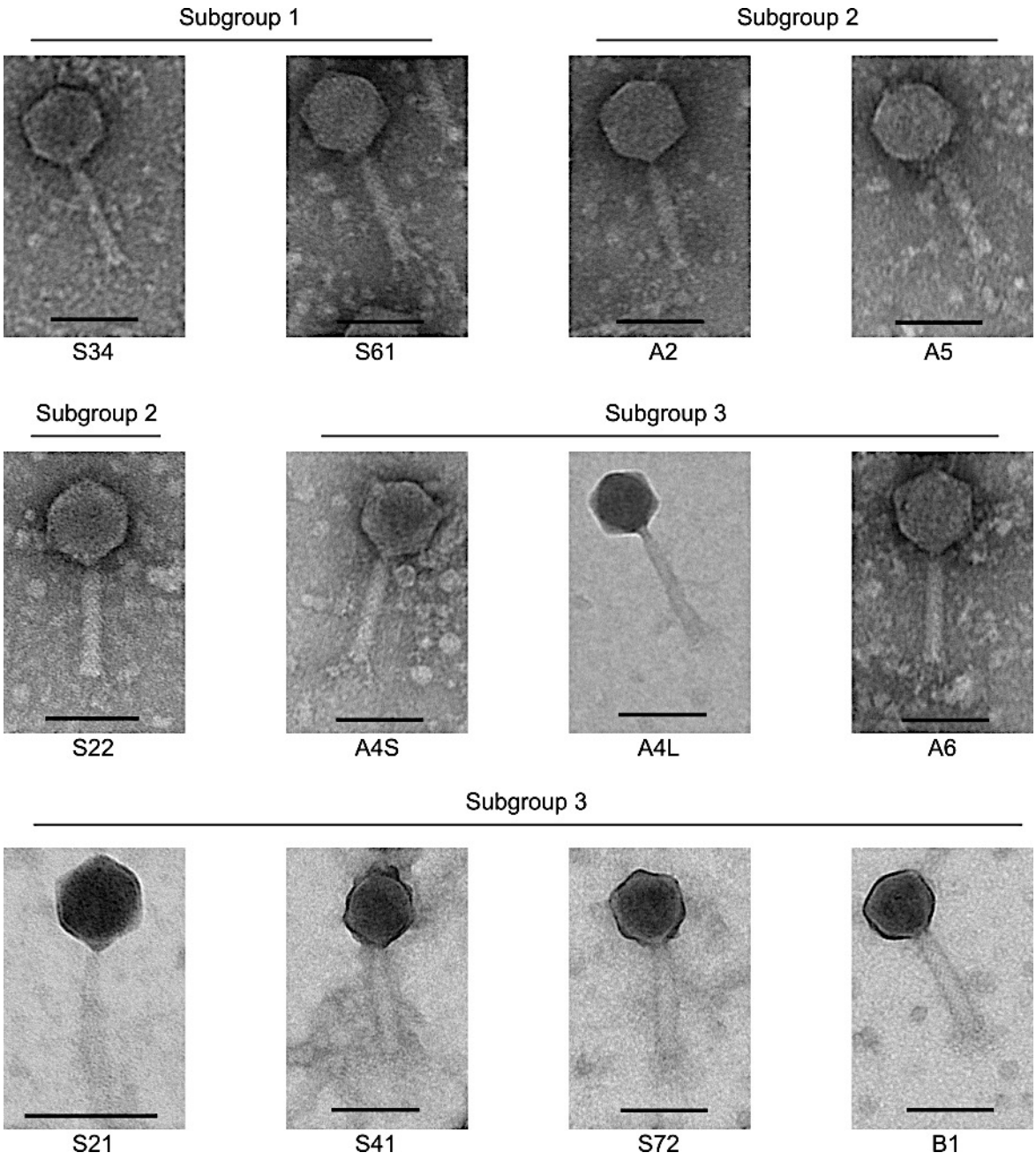



**Figure 3.3** TEM images of individual ΦTE-like phages.

All phages shown belong to the *Myoviridae* family. Black bars for all images represent 100 nm.

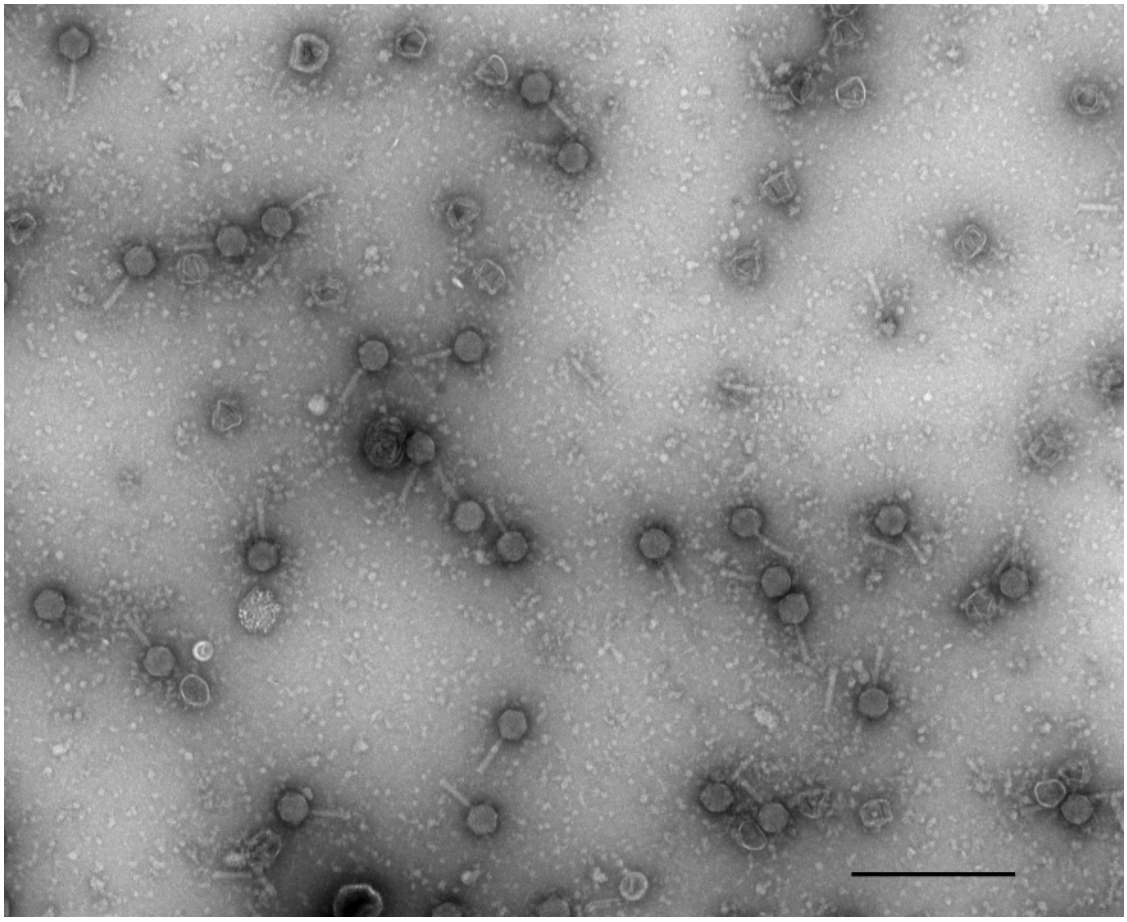


**Figure 3.4**. TEM image of several ΦTE-like phage, ΦA5.
Phages shown in the image all have extended tails. No phages with contracted tails were found. The black bar represents 500 nm.

### 3.2.2 ΦM1-like phages

#### 3.2.2.1 Genomic analysis of ΦM1-like phages

12 of the newly sequenced *P. atrosepticum* phages and two previously submitted phages (Peat1 and PP90) shared high homology with ΦM1. Collectively these phages shared more than 76.3% identity with each other. ΦM1 was classified as a *Phikmvvirus* which belongs to the subfamily *Autographivirinae* (Blower et al., 2017). This subfamily of viruses have linear genomes that are usually flanked by repeating sequences of DNA termed direct terminal repeats (DTRs) (Lavigne et al., 2003). ΦM1's DTR could not be established by primer walking and thus was estimated by analysis of restriction enzyme digestion patterns of its genome (Blower et al., 2017). All ΦM1-like phages were sequenced by Illumina sequencing, thus the short reads often cause the direct repeats to overlap with each other and thus are not present in the final assembled contig. The predicted DTR of ΦM1 was used as a reference to try to predict the DTR regions of the ΦM1-like phages. However, for four of the phages, ΦS33, ΦS42, ΦS71 and ΦB24, DTRs were present in the final assembled contig and thus these could be correctly assigned.

The 12 newly sequenced phages can be split into five further subgroups based on their relatedness. Subgroups in this case shared more than 97.9% identity (**Table 3.3**). ΦRC10 was annotated as a representative of the ΦM1-like phages as sequence comparisons showed that it was the least similar to ΦM1 but it also had an interesting Abi phenotype, discussed in the next chapter. Similar to ΦM1, all ΦRC10 open reading frames are coded on the forward reading frame (**Figure 3.5**). ΦM1-like phages have genome lengths of 43,661 - 45,633 base pairs with 52 to 61 ORFs. Only ΦM1 has a single tRNA, tRNA-Ile, no other ΦM1-like phage annotated contains any tRNAs.

Other non *P. atrosepticum* phages similar to the ΦM1-like phages include two *P. carotovorum* phages, PPWS1 (LC063634, (Hirata et al., 2016)) and PP16 (NC_031068, unpublished) which share 78% and 59% nucleotide identity respectively with ΦRC10. There is also a phage isolated on a *Dickeya* species called BF25/12 (KT240186, unpublished) which shares 48.3% nucleotide identity with ΦRC10. ACT

comparison images generated using tblastx comparison between ΦRC10 and closely related phage genomes shows that their genomes are syntenic (**Figure 3.6**).

| **Table 3.3** Sequence identity between ΦM1-like phages (%) | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Subgroups | | | | | | | | | | | | | | |
| | 1 | | 2 | 3 | | 4 | | | | | | 5 | | | |
| Phage | M1 | B24 | M5 | A1 | M4 | M2 | S32 | S33 | S42 | S64 | S71 | RC9 | RC10 | Peat1 | PP90 |
| M1 | 100 | 99.6 | 90.6 | 88.8 | 87.5 | 86.8 | 87.1 | 87 | 86.7 | 87.2 | 86.9 | 85.2 | 85.2 | 82.2 | 80.9 |
| B24 | 99.6 | 100 | 90.3 | 88.5 | 87.2 | 86.5 | 87 | 87.3 | 87 | 87.1 | 87.2 | 85.1 | 85.1 | 82.3 | 80.6 |
| M5 | 90.6 | 90.3 | 100 | 94.1 | 92.8 | 94.5 | 94.6 | 94.7 | 94.4 | 94.9 | 94.6 | 90.2 | 90.2 | 83.9 | 83.3 |
| A1 | 88.8 | 88.5 | 94.1 | 100 | 97.9 | 94 | 94.3 | 94.3 | 94 | 94.6 | 94.3 | 91.5 | 91.5 | 84 | 81.3 |
| M4 | 87.5 | 87.2 | 92.8 | 97.9 | 100 | 92.6 | 92.8 | 92.9 | 92.6 | 93.1 | 92.8 | 93.1 | 93.1 | 83 | 80.6 |
| M2 | 86.8 | 86.5 | 94.5 | 94 | 92.6 | 100 | 98.9 | 99 | 98.6 | 99.2 | 99.8 | 92.2 | 92.2 | 84.4 | 81.7 |
| S32 | 87.1 | 87 | 94.6 | 94.3 | 92.8 | 98.9 | 100 | 99.5 | 99.1 | 99.7 | 99.3 | 92.6 | 92.6 | 84.9 | 81.8 |
| S33 | 87 | 87.3 | 94.7 | 94.3 | 92.9 | 99 | 99.5 | 100 | 99.6 | 99.8 | 99.9 | 92.6 | 92.6 | 85.1 | 81.7 |
| S42 | 86.7 | 87 | 94.4 | 94 | 92.6 | 98.6 | 99.1 | 99.6 | 100 | 99.4 | 99.7 | 92.4 | 92.4 | 85.1 | 81.4 |
| S64 | 87.2 | 87.1 | 94.9 | 94.6 | 93.1 | 99.2 | 99.7 | 99.8 | 99.4 | 100 | 99.6 | 92.6 | 92.6 | 85.1 | 81.9 |
| S71 | 86.9 | 87.2 | 94.6 | 94.3 | 92.8 | 99.8 | 99.3 | 99.9 | 99.7 | 99.6 | 100 | 92.5 | 92.5 | 85.1 | 81.6 |
| RC9 | 85.2 | 85.1 | 90.2 | 91.5 | 93.1 | 92.2 | 92.6 | 92.6 | 92.4 | 92.6 | 92.5 | 100 | 100 | 83 | 82.4 |
| RC10 | 85.2 | 85.1 | 90.2 | 91.5 | 93.1 | 92.2 | 92.6 | 92.6 | 92.4 | 92.6 | 92.5 | 100 | 100 | 83 | 82.4 |
| Peat1 | 82.2 | 82.3 | 83.9 | 84 | 83 | 84.4 | 84.9 | 85.1 | 85.1 | 85.1 | 85.1 | 83 | 83 | 100 | 76.3 |
| PP90 | 80.9 | 80.6 | 83.3 | 81.3 | 80.6 | 81.7 | 81.8 | 81.7 | 81.4 | 81.9 | 81.6 | 82.4 | 82.4 | 76.3 | 100 |

Colours produced by conditional formatting on Microsoft Excel as with **Table 3.1**.

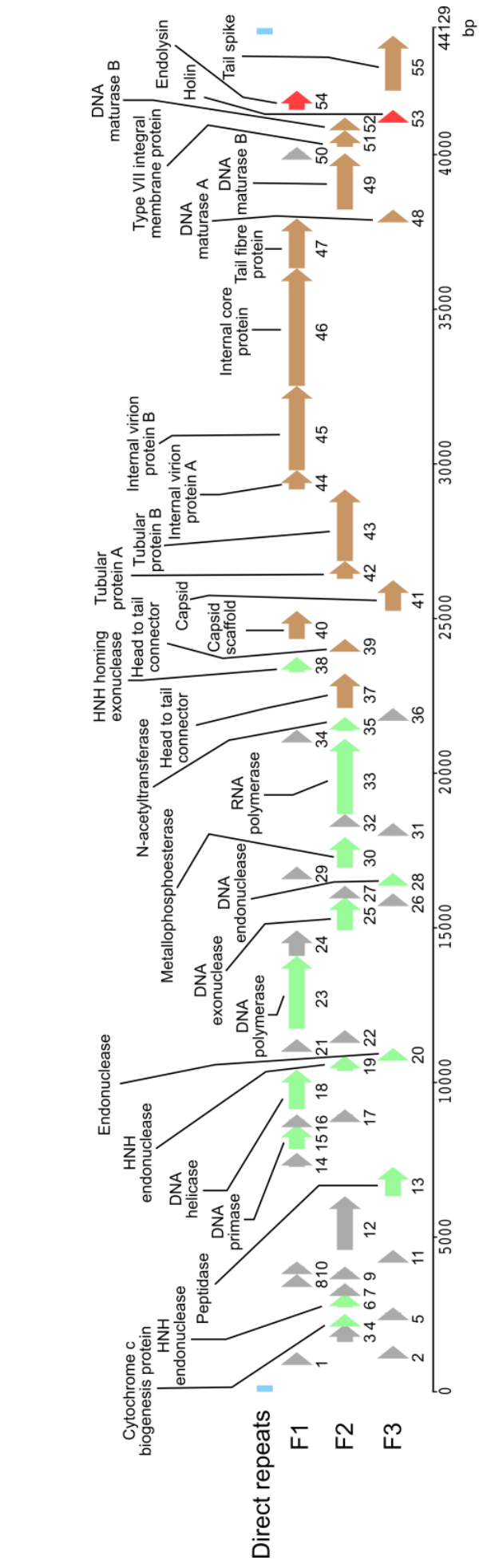



**Figure 3.5** Genomic map of ΦRC10.
All predicted ORFs of ΦRC10 are encoded on the forward reading strand. The different reading frames are called F1-F3. Green, brown and red arrows encode proteins predicted to be involved in metabolism, virion structure and lysis, respectively. Grey arrows represents ORFs that have unknown functions. Blue bars represent direct terminal repeats that are found on both sides of the genome.

**Table 3.4** Genomic data of ΦM1-like phages

| Phage | Subgroup | Genome size (bp) | Predicted ORFs | tRNAs |
|---|---|---|---|---|
| M1 | 1 | 43,827 | 52 | tRNA-Ile |
| B24 | 1 | 43,661 | NA | NA |
| M5 | 2 | 43,984 | NA | NA |
| A1 | 3 | 43,726 | NA | NA |
| M4 | 3 | 44,418 | NA | NA |
| M2 | 4 | 44,487 | NA | NA |
| S32 | 4 | 44,235 | NA | NA |
| S33 | 4 | 44,024 | NA | NA |
| S42 | 4 | 43,898 | NA | NA |
| S64 | 4 | 44,132 | NA | NA |
| S71 | 4 | 43,960 | NA | NA |
| RC9 | 5 | 44,125 | NA | NA |
| RC10 | 5 | 44,129 | 55 | None |
| Peat1 | - | 45,633 | 61 | None |
| PP90 | - | 44,570 | 56 | None |

NA – Not annotated

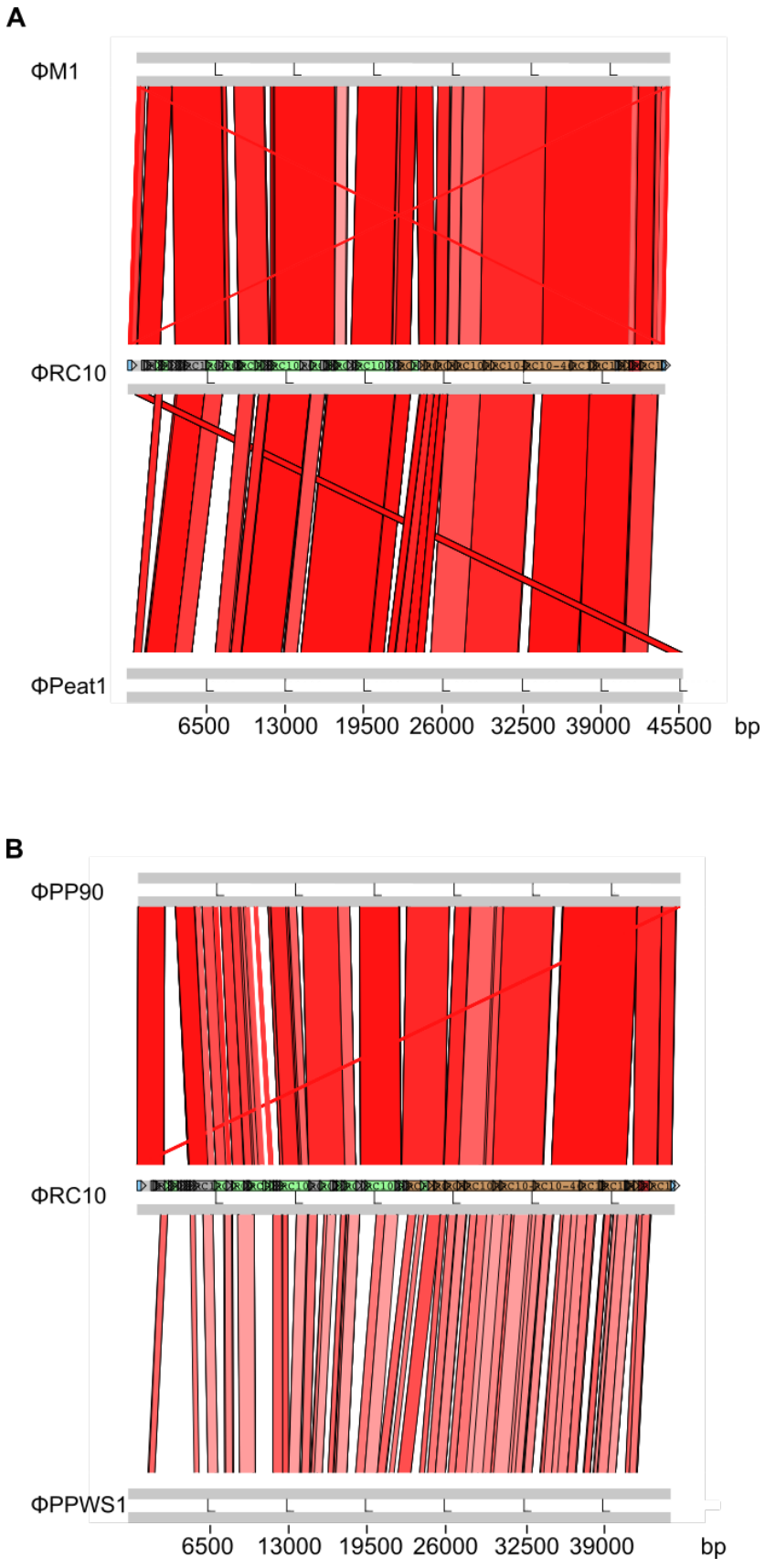



**Figure 3.6** continued overleaf

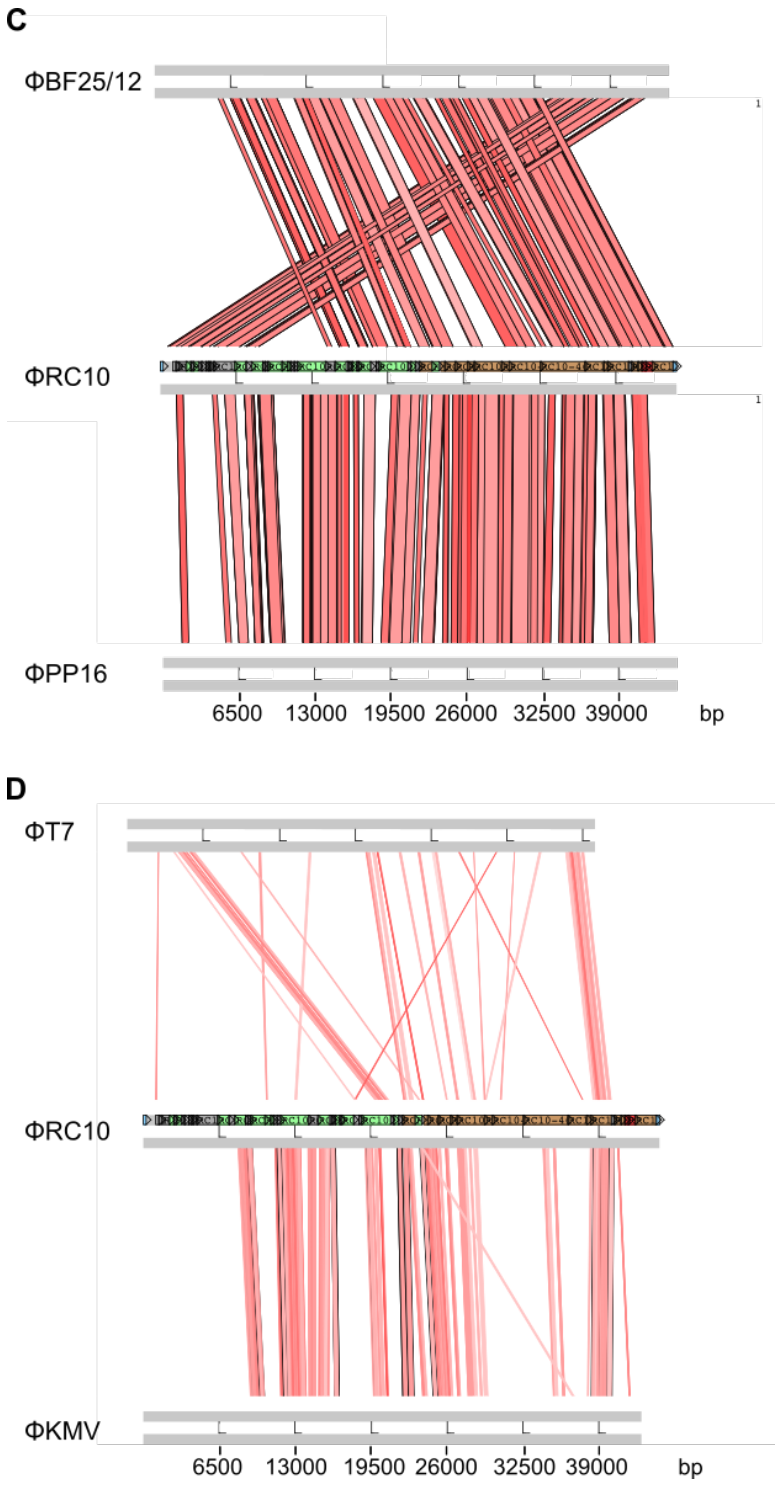



**Figure 3.6** Genome wide protein comparison of ΦRC10 and related phages. **A**. **Φ**RC10 compared with ΦM1 and Peat1. **B**. ΦRC10 compared with *P. carotovorum* phages, PP90 and PPWS1. **C**. ΦRC10 compared to *P. carotovorum* phage ΦPP16 and *Dickeya* phage ΦBF25/12. **D**. ΦRC10 compared to ΦT7 and ΦKMV. Comparisons files were made using tblastx analysis with a cut off value of $1x10^{-4}$. Red bars signify areas of high amino acid sequence homology.

### 3.2.2.2 TEM analysis of ΦM1-like phages

As expected from the genomic data, all ΦM1-like phages were proven to be members of the *Podoviridae* family (**Figure 3.7**). This family is defined as double stranded DNA phages with a short non contracting tail. Heads measure from 57 nm to 69 nm which is slightly larger than the ΦM1 original measurements of 55 nm (Toth et al., 1997).

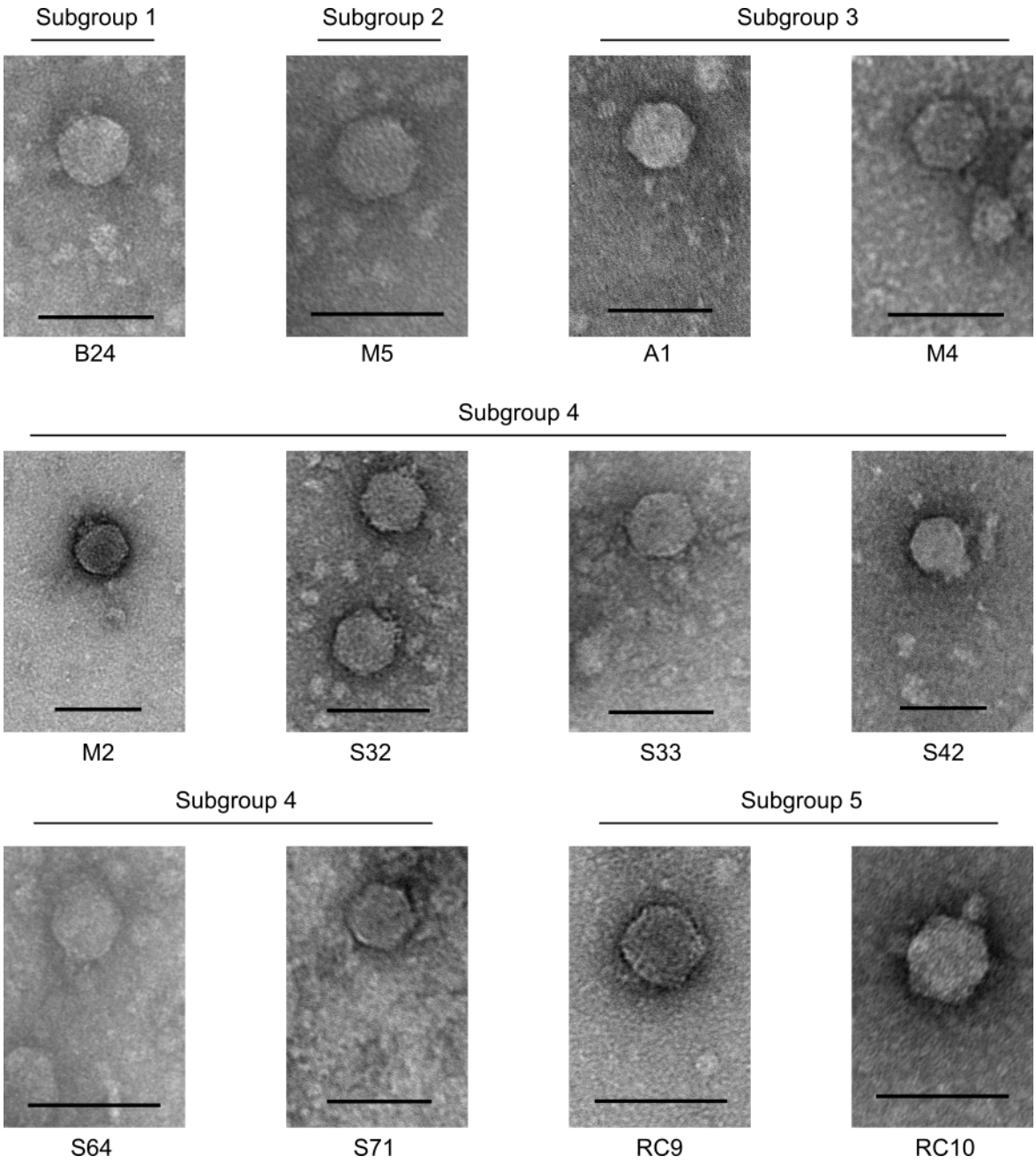



**Figure 3.7** TEM images of ΦM1-like phages.
All phages belong to the *Podoviridae* family. Black bars represent 100 nm.

### 3.2.3 ΦR1-like phages

#### 3.2.3.1 Genomic analysis of ΦR1-like phages

ΦR1-like phages are a group of eight phages from the library with genomes that displayed more than 93% identity to each other (**Table 3.5**). As there were no *P. atrosepticum* phages sequenced that resemble this type of phage as of yet, they were named after the first phage isolated of the group called ΦR1. Blastn analysis of the genomes revealed that they have homology with podoviruses from the *T7virus* genus. This genus is in the subfamily *Autographivirinae*, the same as the ΦM1-like phages. Similar to ΦM1-like phages the ΦR1-like phages were predicted to have linear genomes flanked by DTRs. The location of the DTRs of the ΦR1-like phages are based on two similar T7viruses that have their DTRs annotated in the NCBI database, *Yersinia* phage vB YenP AP10 (direct submission, AP10 for short, KT852574.1) and *Citrobacter* phage SH2 (direct submission, KU687348.1). Addition of the DTRs show that the lengths of the ΦR1-like phages range from 40,255 to 41,375 (**Table 3.6**).

As ΦR1 was the establishing member of this group it was chosen for annotation. This revealed that, similar to the ΦM1-like phages, all 50 predicted open reading frames of ΦR1 are coded on the forward reading frame (**Figure 3.8**). There were no tRNA genes detected in the genome. Nucleotide comparisons of ΦR1 showed that it shared 59.7%, 58.8%, 62.1% and 60.5% nucleotide identity with T3 (KC960671.1), T7 (GU071091.1), AP10 and SH2 respectively. ΦR1 has very high nucleotide sequence identity, (96%), with a previously submitted *P. carotovorum* phage, Jarilo (direct submission, MH059637.1). This percentage identity is even higher than certain ΦR1-like phages when compared to each other (**Table 3.5**). Compared to ΦM1 and KMV, ΦR1 shows much less homology sharing sequence identities of 48% with both phages. ACT generated images of ΦR1 and other *T7viruses* show that there are no large rearrangements between the phage genomes of this genus. (**Figure 3.9, A-C**). Compared to ΦM1 and ΦKMV, however, there is much less amino acid homology seen in the predicted products (**Figure 3.9, D**).

| **Table 3.5** Sequence identity between ΦR1-like phages | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Phages | | | | | | | | |
| | ΦR1 | ΦRC58 | ΦRC64 | ΦRC41 | ΦRC30 | ΦRC36 | ΦRC33 | ΦRC18 |
| ΦR1 | 100 | 97.2 | 96.5 | 96.1 | 94.9 | 95.3 | 95.2 | 93.7 |
| ΦRC58 | 97.2 | 100 | 98 | 97 | 96.6 | 96.9 | 96.8 | 95.3 |
| ΦRC64 | 96.5 | 98 | 100 | 97.2 | 96.1 | 96.5 | 96.2 | 94.9 |
| ΦRC41 | 96.1 | 97 | 97.2 | 100 | 96.7 | 97.2 | 95.2 | 95.6 |
| ΦRC30 | 94.9 | 96.6 | 96.1 | 96.7 | 100 | 98.9 | 94.6 | 97 |
| ΦRC36 | 95.3 | 96.9 | 96.5 | 97.2 | 98.9 | 100 | 95.1 | 97.5 |
| ΦRC33 | 95.2 | 96.8 | 96.2 | 95.2 | 94.6 | 95.1 | 100 | 93.5 |
| ΦRC18 | 93.7 | 95.3 | 94.9 | 95.6 | 97 | 97.5 | 93.5 | 100 |

Colours produced by conditional formatting on Microsoft Excel as with **Table 3.1**.

**Table 3.6** Genomic data of ΦR1-like phages

| Phage | Genome size (bp) | Predicted ORFs |
|---|---|---|
| ΦR1 | 40,728 | 50 |
| ΦRC58 | 40,307 | NA |
| ΦRC64 | 40,255 | NA |
| ΦRC41 | 40,686 | NA |
| ΦRC30 | 41,375 | NA |
| ΦRC36 | 41,233 | NA |
| ΦRC33* | 39,494 | NA |
| ΦRC18 | 40,553 | NA |

*ΦRC33's raw sequence data returned in three contigs and so is not a complete genome and will need to be resolved by resequencing the gaps.

NA – not annotated

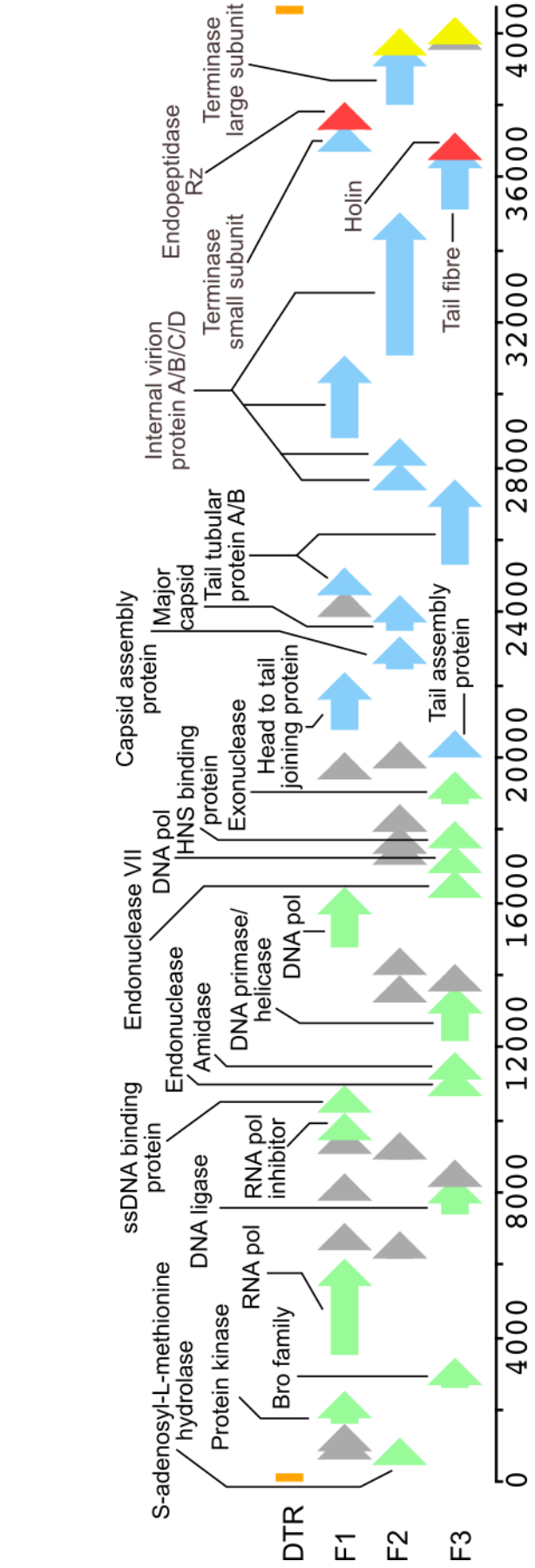



**Figure 3.8** Genomic map of ΦR1.

All predicted open reading frames are encoded on the forward reading strand. F1-F3 indicates the reading frame. Green, blue and red arrows represent ORFs involved in metabolism, virion structure and lysis, respectively. Grey arrows represent ORFs that have homology with other phage proteins with no known functions. Yellow arrows represent encoded proteins that do not have any significant hits on NCBI. The orange bars on either end of the genome represent the direct terminal repeats.

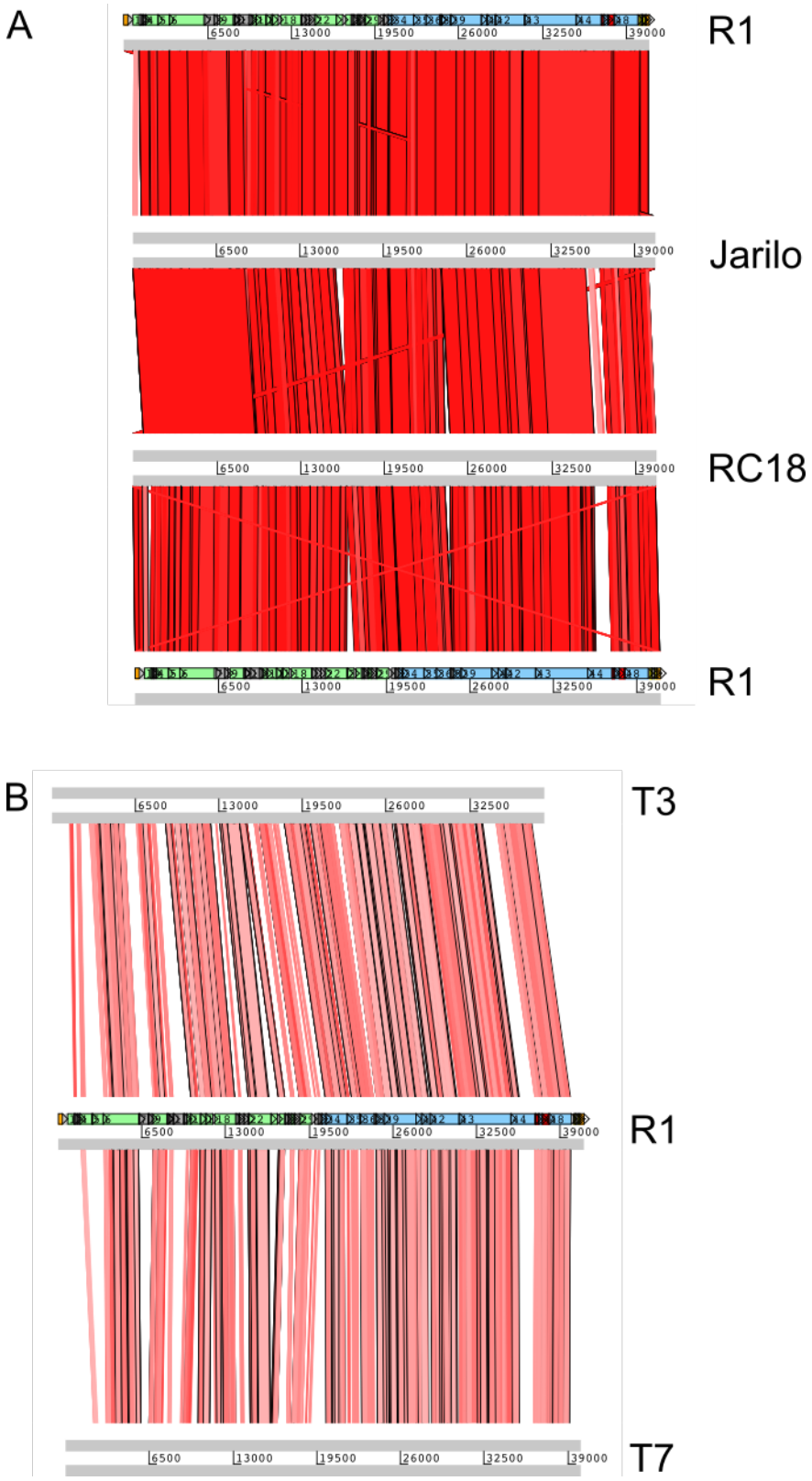



**Figure 3.9** continued overleaf

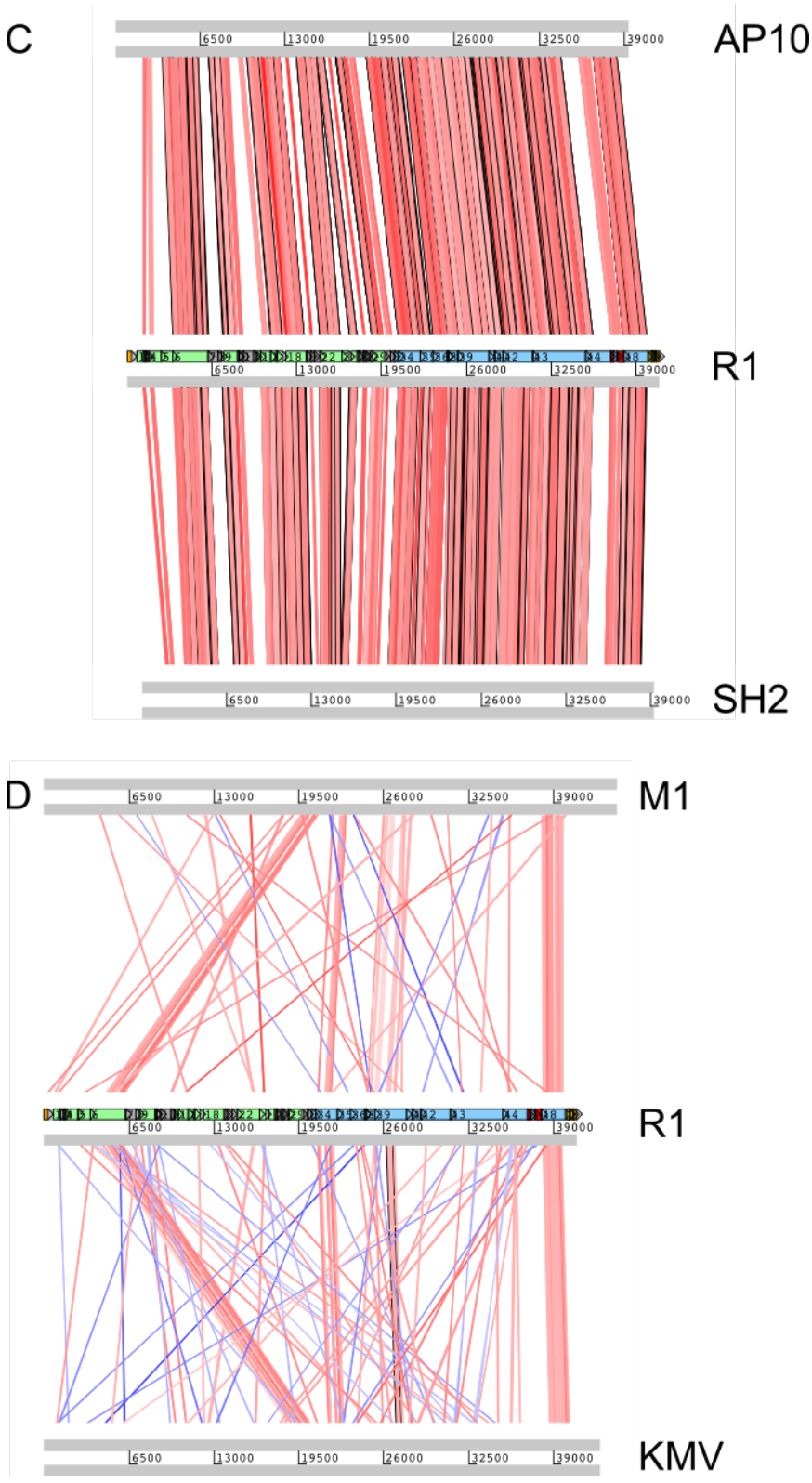



**Figure 3.9** Genome wide predicted protein comparisons of ΦR1 and related phages. **A**. ΦR1 compared with closely related *P. carotovorum* phage Jarilo and another ΦR1-like phage, ΦRC18. **B**. ΦR1 compared with T-series phages, T3 and T7. **C**. ΦR1

compared to *Yersinia* phage AP10 and *Citrobacter* phage SH2. **D**. ΦR1 compared to ΦM1 and KMV.

### 3.2.3.2 TEM analysis of ΦR1-like phages

As expected all ΦR1-like phages when viewed under TEM resemble podoviruses. Their heads measure 56 to 68 nm in diameter which is of similar sizes to ΦM1-like particles.

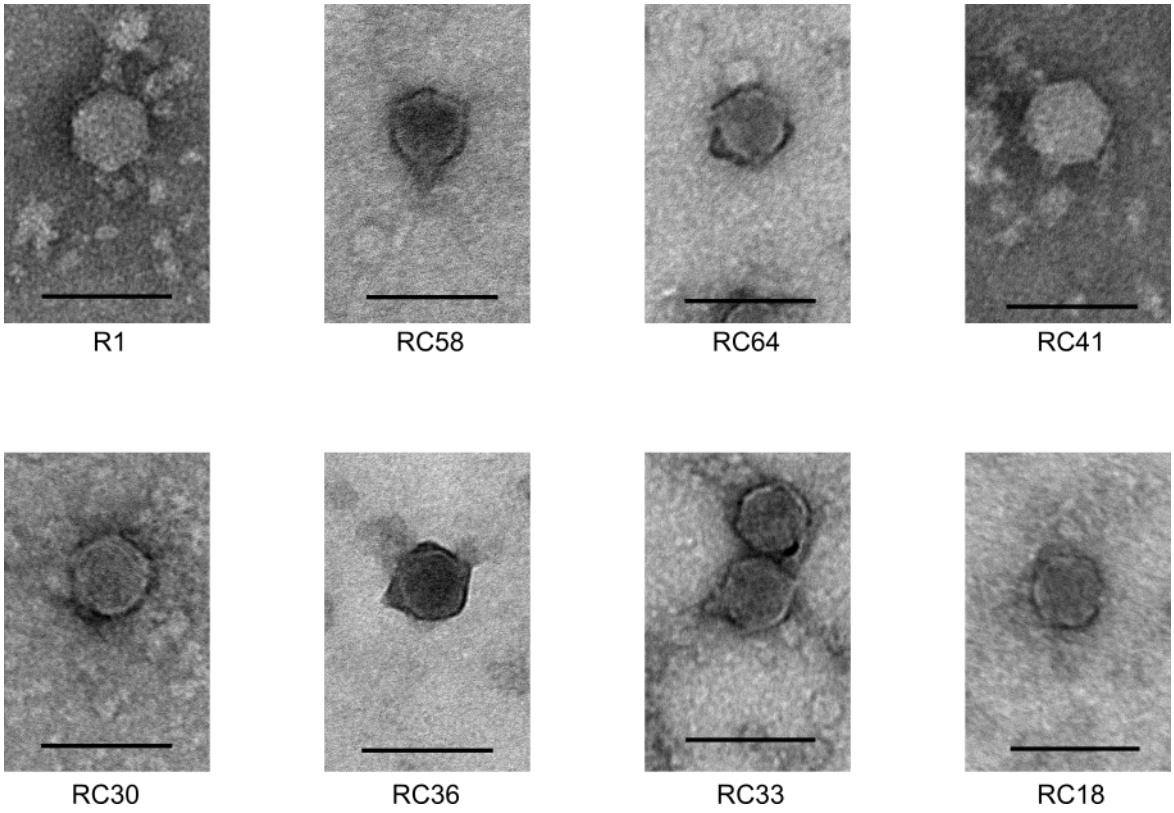



**Figure 3.9** TEM images of ΦR1-like phages.

ΦR1-like phages all belong to the *Podoviridae* family. Black bars represent 100 nm.

### 3.2.4 ΦRC6-like phages

#### 3.2.4.1 Genomic analysis of ΦRC6-like phages

The ΦRC6-like phages are a set of three phages that share more than 94.5% sequence identity to each other (**Table 3.7**). Blastn results of the ΦRC6-like phages show very poor homology with phages belonging to the *Ounavirinae* subfamily. This subfamily is currently comprised of four genera which are the *Ea214virus*, *Felix01virus*, *Mooglevirus* and *Suspvirus* named after the first phages isolated for each group. Comparisons of the genomes of the founding phages of each genus show that the ΦRC6-like viruses do not closely match any of the genera and taxonomically should define a new genus (**Table 3.7**).

As the founding member of the ΦRC6-like phages, ΦRC6 was annotated. A genomic map of ΦRC6 is shown in **Figure 3.9**. The ΦRC6-like phages have genomes that range from 93,058 to 93,203 base pairs. Annotation of ΦRC6 has revealed that it has 151 open reading frames with 24 tRNAs (**Table 3.8**). The large number of tRNAs is a common feature of the *Ounavirinae* with the founding members of each genus carrying 19 to 26 tRNAs (**Table 3.8**). Another interesting feature of ΦRC6 is that it contains a possible CRISPR repeat within its genome (**Figure 3.10**). This putative CRISPR repeat contains three repeats of 33 base pairs with spacer lengths of 56 and 61 base pairs. CRISPR repeats also are preceded by a 100 to 500 base pair long AT rich leader containing a promoter (Wei et al., 2015). The putative leader is shown in **Figure 3.10B** beginning at the predicted -35 promoter region and has an AT content of 66.1%. However, it appears that the -10 promoter region within the leader is not essential for the acquisition of new spacers (Yosef et al., 2012).

**Table 3.7.** Nucleotide identities of ΦRC6-like viruses and *Ounavirinae*

| Phages | ΦRC6 | ΦRC50 | ΦRC62 | Ea214 | Felix01 | Moogle | SUSP1 |
|---|---|---|---|---|---|---|---|
| ΦRC6 | 100 | 94.4 | 94.5 | 48.9 | 50.6 | 50.2 | 51 |
| ΦRC50 | 94.4 | 100 | 99.9 | 49.2 | 50.6 | 50.2 | 51 |
| ΦRC62 | 94.5 | 99.9 | 100 | 49.2 | 50.6 | 50.2 | 51.1 |
| Ea214 | 48.9 | 49.2 | 49.2 | 100 | 51.3 | 51.9 | 51.6 |
| Felix01 | 50.6 | 50.6 | 50.6 | 51.3 | 100 | 70.5 | 71.1 |
| Moogle | 50.2 | 50.2 | 50.2 | 51.9 | 70.5 | 100 | 79.2 |
| SUSP1 | 51 | 51 | 51.1 | 51.6 | 71.1 | 79.2 | 100 |

Colours produced by conditional formatting on Microsoft Excel as with **Table 3.1**.

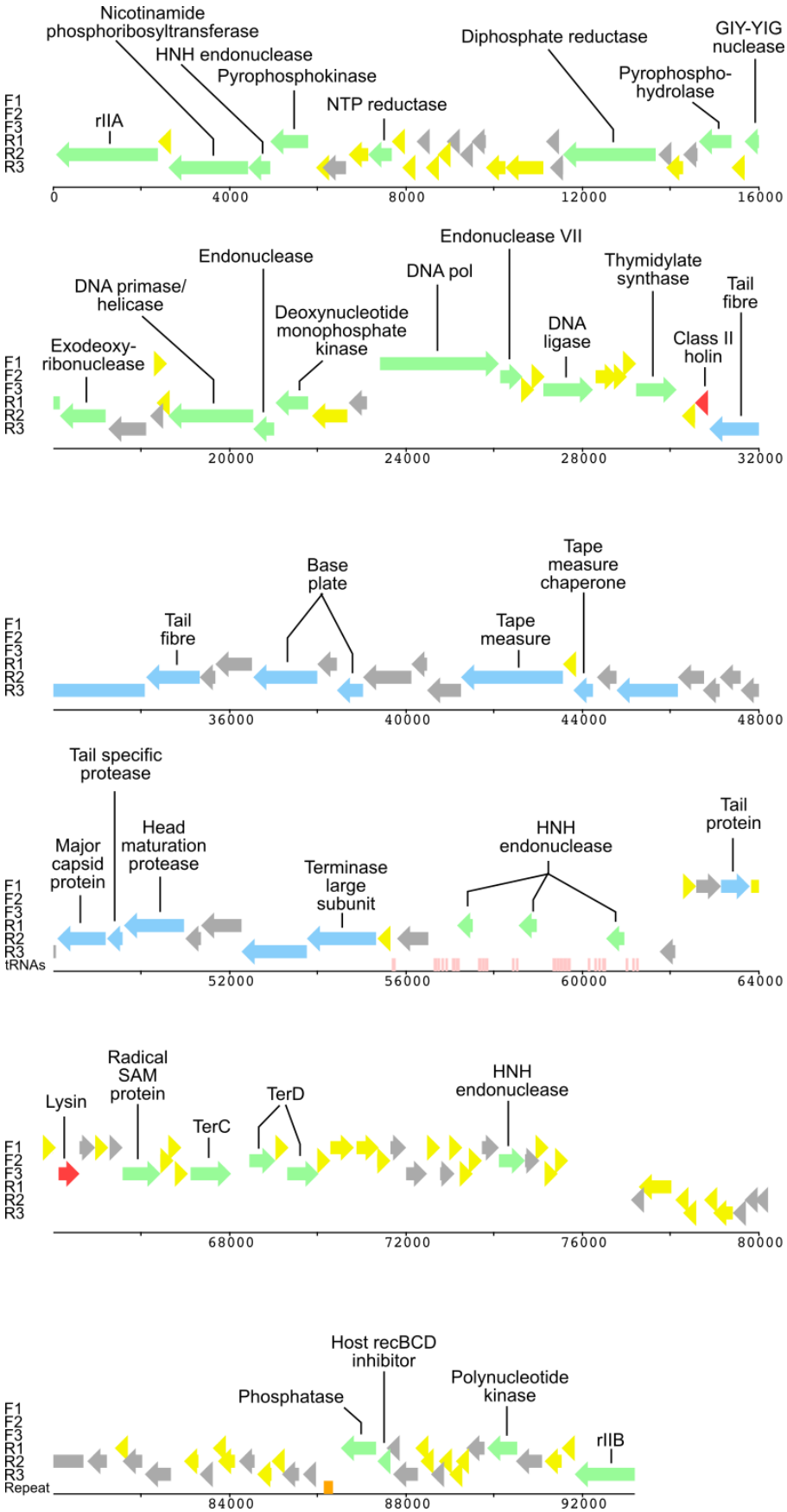



**Figure legend shown overleaf**

**Figure 3.9** Genomic map of ΦRC6.

F1-3 indicates ORFs on the forward reading strand with R1-3 for the reverse. Arrows in green, blue and red are ORFs involved in metabolism, virion structure and lysis, repectively. ORFs encoding proteins of unknown functions that have homologs in other phages are in grey. Yellow arrows represent ORFs that have no significant matches in NCBI. Pink bars represent the location of tRNAs with the putative CRISPR repeat displayed as an orange bar.

**Table 3.8** Genomic data of ΦRC6-like phages

| Phage | Genome length (bp) | ORFs | tRNAs |
|---|---|---|---|
| ΦRC6 | 93,203 | 151 | 24 |
| ΦRC50 | 93,185 | NA | NA |
| ΦRC62 | 93,058 | NA | NA |
| Ea214 | 84,576 | 118 | 26 |
| Felix01 | 86,155 | 247 | 22 |
| Moogle | 87,999 | 126 | 21 |
| SUSP1 | 90,743 | 138 | 19 |

NA – not annotated

**A**


**B**

| Component | Sequence |
|---|---|
| Leader | **TTTAGA**AATATTGGGCATTTTTC**TTTTAAAAT**CT TCTTCAAAATTCCCCGCAAGATACCTTATGAAG TACCTCTTTACATATCAAACACTTAAACTACACC CTTCATAACGTCCTCCTAAAAATCTA TCCAGAAAACTGTG |
| Repeat | TCACATTGCAAAACTCGTTGGAAAAAAGTGCTC |
| Spacer 1 | TAGGTAGTTATGCATAAAATTTCCTGTGCA GTTTACTTCACATTGGCTCACAAATA |
| Spacer 2 | CCTAGACGTGAATGTCACATTGCAAA ACTCACTGCAAAAAAGTGCTCTCAATTCGCCAGTG |

**Figure 3.10** ΦRC6 contains a putative CRISPR repeat.

**A** Organisation of the putative CRISPR. The putative CRISPR repeat in ΦRC6 contains an AT rich leader sequence followed by three 36 bp direct repeats, interspersed by two different spacer regions. The sequences of the different components of the putative CRISPR repeat is shown in **B**. Sequences of the different regions. Letters in bold in the leader sequence are the predicted -35 (TTTAGA) and -10 (TTTTAAAAT) sequences.

### 3.2.4.2 TEM analysis of ΦRC6-like phages

Phages of the *Ounavirinae* subfamily are part of the *Myoviridae* family. The ΦRC6-like phages were confirmed to be Myoviruses through TEM analysis (**Figure 3.11**). The heads of ΦRC6-like phages measure 57 to 68 nm across with an extended tail of 115 to 121 nm.

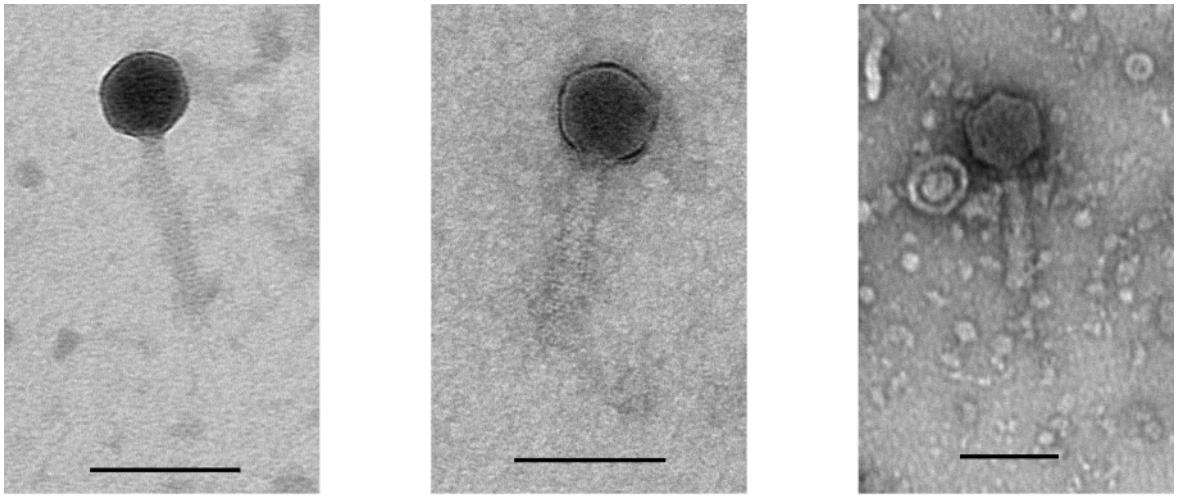



**Figure 3.11** TEM images of ΦRC6-like phages.
All phages observed belong to the *Myoviridae* family. Black bars represent 100 nm.

### 3.2.5 ΦRC63-like phages

#### 3.2.5.1 Genomic analysis of ΦRC63-like phages

The ΦRC63-like phage "group" currently contains only a single phage, ΦRC63. Analysis of the genome shows that it can be classed as a jumbo phage with a genome size of 240, 968 base pairs long. Jumbo phages have genomes over 200, 000 base pairs long and are considered relatively rare phages (Yuan and Gao, 2017). Annotation of the ΦRC63 genome revealed that it is circular with 271 predicted ORFs (**Figure 3.12**). Blastn results of the ΦRC63 genome leads to no phages hits. However, during annotation of the genome several proteins showed homology to other submitted jumbo phage genomes. The most common significant hit came from proteins originating from an *Erwinia amylovora* phage called Risingsun (MF_49646). Other related phages include *Pseudomonas aeruginosa* phage EI (NC_007623), *E. amylovora* phage Caitlin (NC_031120) and *Salmonella* phage SPN3U5 (NC_027402). All these phages are jumbo phages but share very low nucleotide sequence homology with each other (**Table 3.9**). ACT comparison images using blastn and tblastx comparisons show that amino acid homology is much more conserved than nucleotide homology suggesting convergent evolution (**Figure 3.13**). Like many of the other jumbo phages, many of the predicted open reading frames have no known function. For ΦRC63, 193 of the ORFs are unknown which accounts for 71.2% of the open reading frames. Another interesting feature to note is that while both the forward and reverse strands encode predicted ORFs, there are only 34 predicted ORFs on the reverse strand with the majority of ORFs coded on the forward reading strand. The location of the ORFs on the reverse strand are also dotted around the genome rather than in localised areas (**Figure 3.12**).

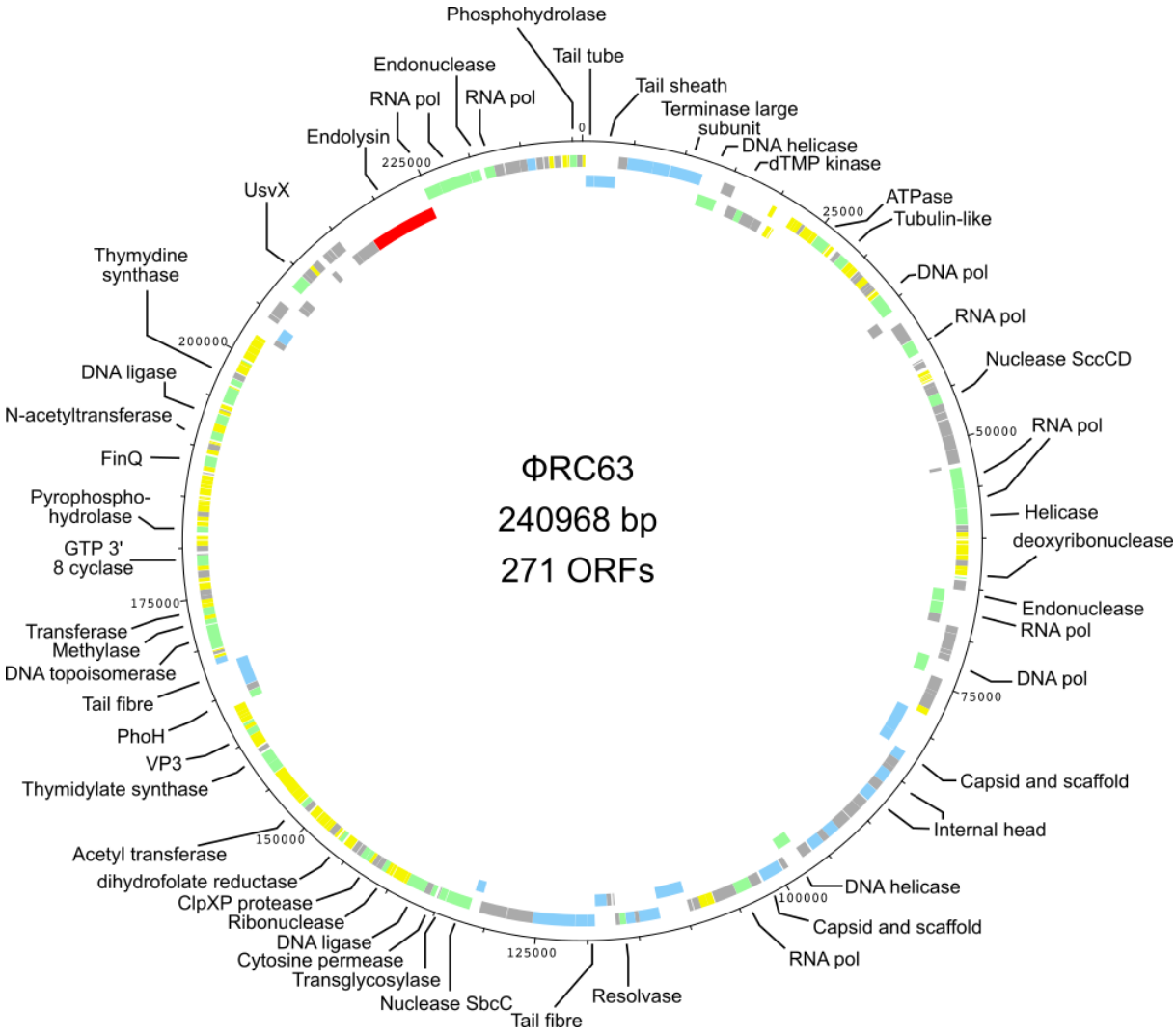



**Figure 3.12** Genomic map of ΦRC63.

Coloured boxes on the outer ring represent ORFs on the forward reading frame with boxes on the inner ring representing ORFs on the reverse reading frame. Green, blue and red boxes represent ORFs encoding proteins involved in metabolism, virion structure and lysis, repectively. Grey boxes represent ORFs encoding proteins of an unknown function that have homologs in other phages. Yellow boxes represent ORFs encoding proteins that have no homologous proteins in NCBI.

**Table 3.9** Nucleotide comparisons of ΦRC63 and related phages

| | ΦRC63 | Rising sun | El | Caitlin | SPN3U5 |
|---|---|---|---|---|---|
| ΦRC63 | 100 | 47.4 | 46.9 | 46.2 | 47 |
| Rising sun | 47.4 | 100 | 48.1 | 47.7 | 47.8 |
| El | 46.9 | 48.1 | 100 | 47.1 | 46.9 |
| Caitlin | 46.2 | 47.7 | 47.1 | 100 | 56.9 |
| SPN3U5 | 47 | 47.8 | 46.9 | 56.9 | 100 |

Colours produced by conditional formatting on Microsoft Excel as with **Table 3.1**.

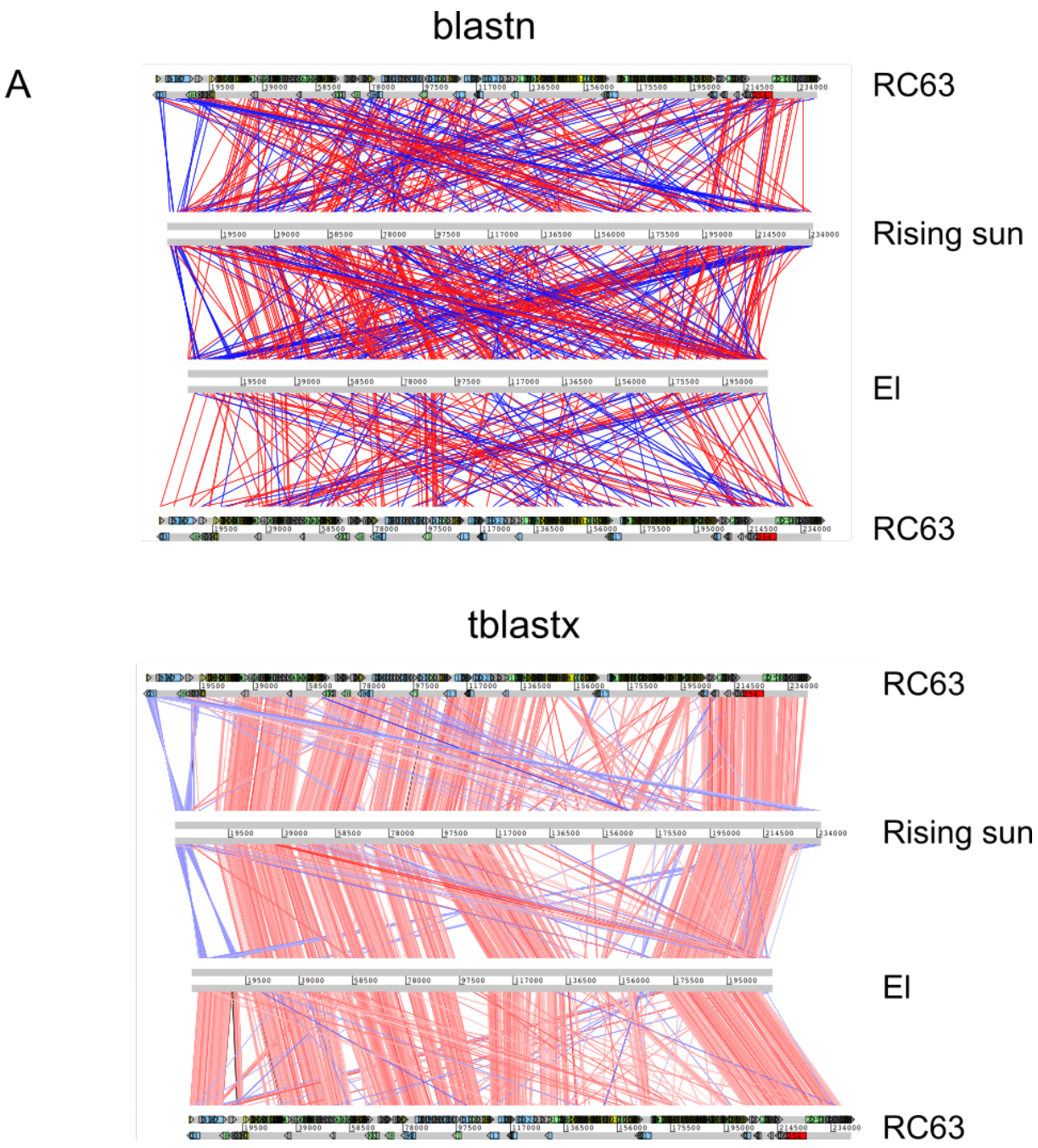



**Figure 3.13** continued overleaf

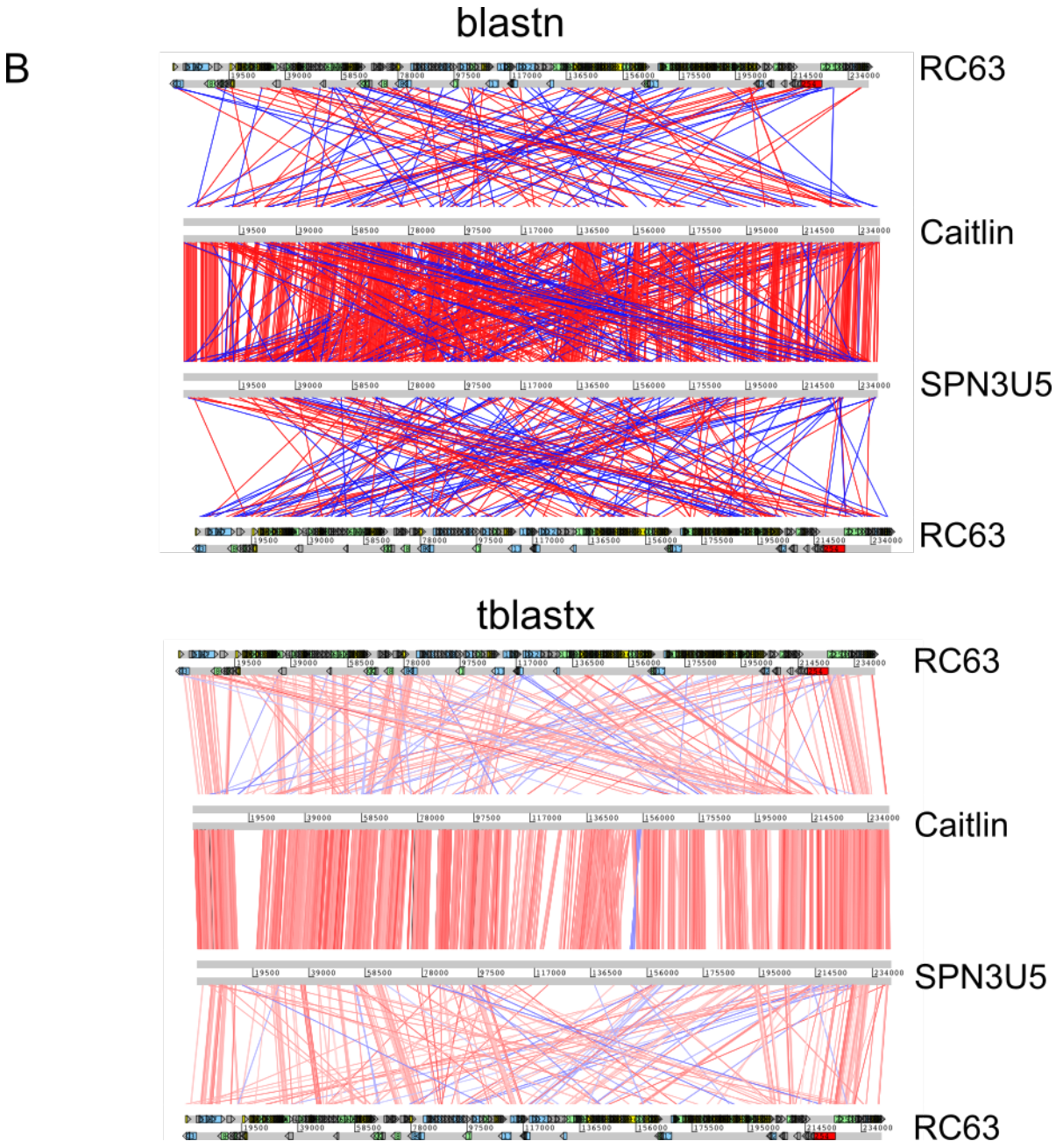



**Figure 3.13** Genome comparisons of ΦRC63 and related phages.

**A**. Blastn and tblastx comparisons of ΦRC63 with *E. amylovora* phage Risingsun and *P. aeruginosa* phage El. **B**. Blastn and tblastx comparisons of ΦRC63 with *E. amylovora* phage Caitlin and *Salmonella* phage SPN3U5.

### 3.2.5.2 TEM analysis of ΦRC63-like phage

High titre purified lysates of ΦRC63 were prepared for TEM. Images obtained confirmed that it was a member of the *Myoviridae*. The size of a ΦRC63 particle is larger than previous *Myoviridae* phages of *P. atrosepticum*. The head measures 118 nm across with an extended tail of 143 nm. This is approximately 1.5 times larger than the head of ΦTE. The larger size of the particle is presumably a result of harbouring a large genome.

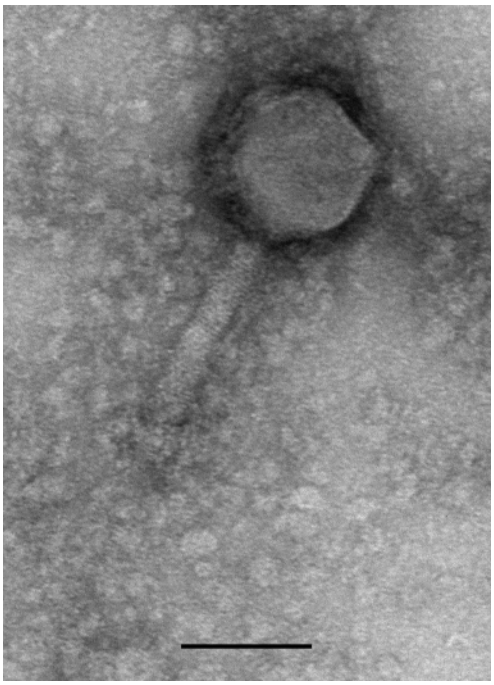


**Figure 3.14** TEM image of ΦRC63**.**
Black bar represents 100 nm.

### 3.2.6 ΦR11-like phages

#### 3.2.6.1 Genomic analysis of ΦR11-like phages

The final "group" of *P. atrosepticum* phages is the ΦR11-like group that currently comprises of only one member, ΦR11. Like ΦRC63, ΦR11 is also a jumbo phage possessing an even larger genome of 269,692 base pairs (**Figure 3.15**). It encodes 354 ORFs which share homology with multiple ORFs from two *E. amylovora* jumbo phages, Y3 (KY984068) and Yoloswag (KY448244). Of the 354 predicted ORFs, 293 have no known function, and 161 of the 293 ORFs do not encode products exhibiting homology with any proteins. As with Y3 and Yoloswag, ΦR11 does not encode any tRNAs. Other related phages include *P. aeruginosa* phage, PaBG (NC022096.1), *P. putida* phage, Lu11 (NC017972.1), NTCB and *Ralstonia* phage RSL1 (NC010811.2). There is little information on the other phages as they represent a very rare class of jumbo phages that have been recently sequenced with very few of their proteins sharing homology with any other phage apart from each other.

Due to its similarity to Y3 and Yoloswag, the genome of ΦR11 has been orientated to start at the same start site which encodes a putative helicase. Homologues of this helicase are also found in the other related jumbo phages mentioned above and so their genomes have been reoriented to begin at this helicase gene for genome comparisons. ACT generated tblastx images show that the genes encoding the products of these phages compared to ΦR11 are disordered (**Figure 3.16**). Genes encoding many of the homologous proteins begin in different areas from the expected positions based on ΦR11. One example of this is a terminase in ΦR11 located at position 2522 to 5329. For Y3 (ARW58644.1) and Yoloswag (AQT28491.1) the homologs are located in similar locations as well, 2538 to 5369 and 2470 to 5283, respectively. However, the homolog in RSL1 (YP001960095.1) is located in a completely different location position, 166,748 to 168,940 (location based on RSL1 genome beginning with the same helicase as ΦR11). Many other regions of the phage are also inverted suggesting that these phages have plastic genomes with many genes capable of rearrangement to different areas of the genome.

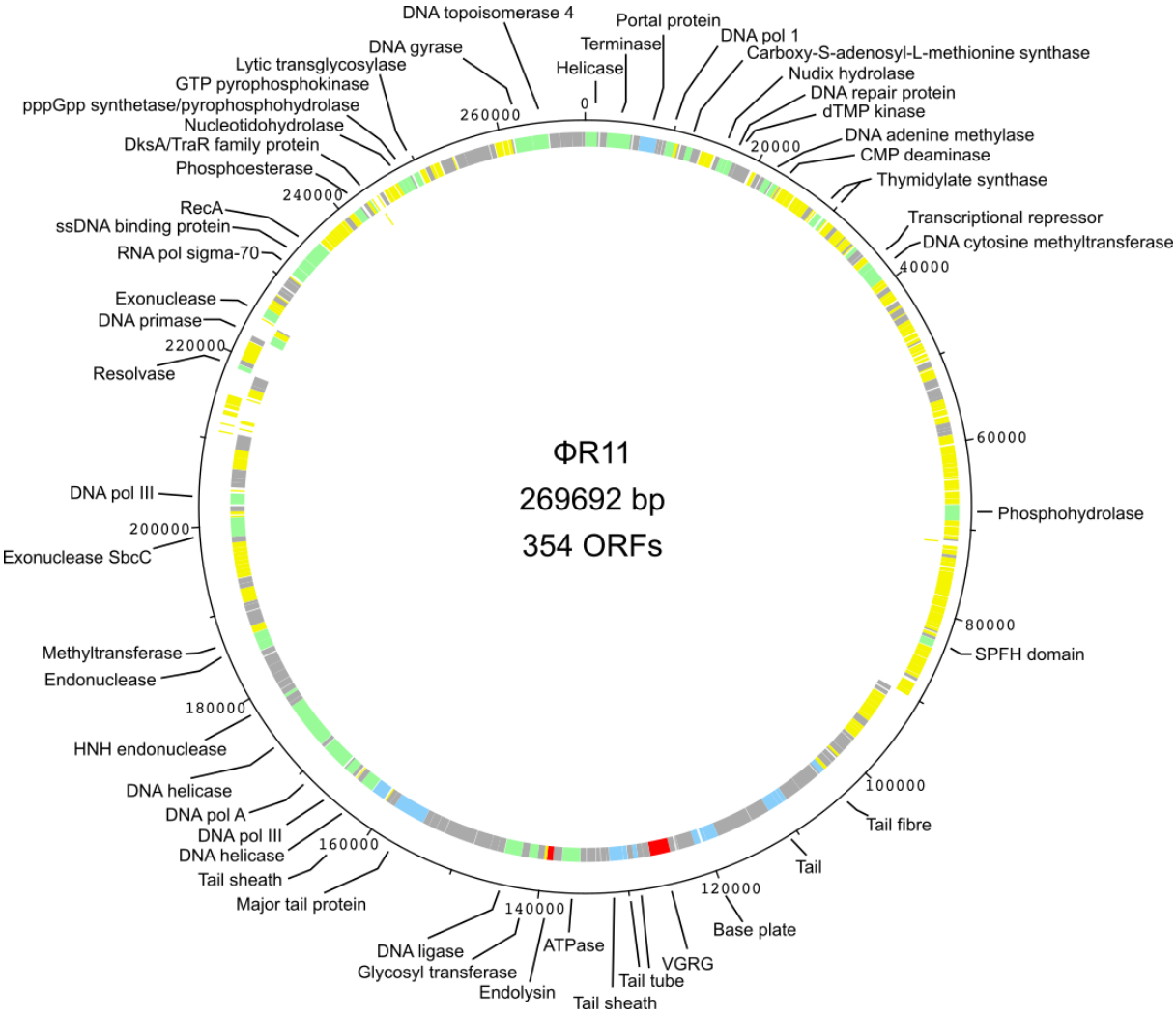



**Figure 3.15** Genomic map of ΦR11.

ΦR11 has a circular double stranded DNA genome of 269, 692 base pairs encoding 354 predicted ORFs. Green, blue and red boxes, indicate genes involved in metabolism, virion structure and lysis, respectively. Grey boxes represents genes of unknown functions with homologs in other phages. Yellow boxes indicate genes with unknown functions with no significant hits on the NCBI database. Boxes on the outer ring are ORFs encoded in the forward reading strand with boxes on the inner ring representing those on the reverse strand.

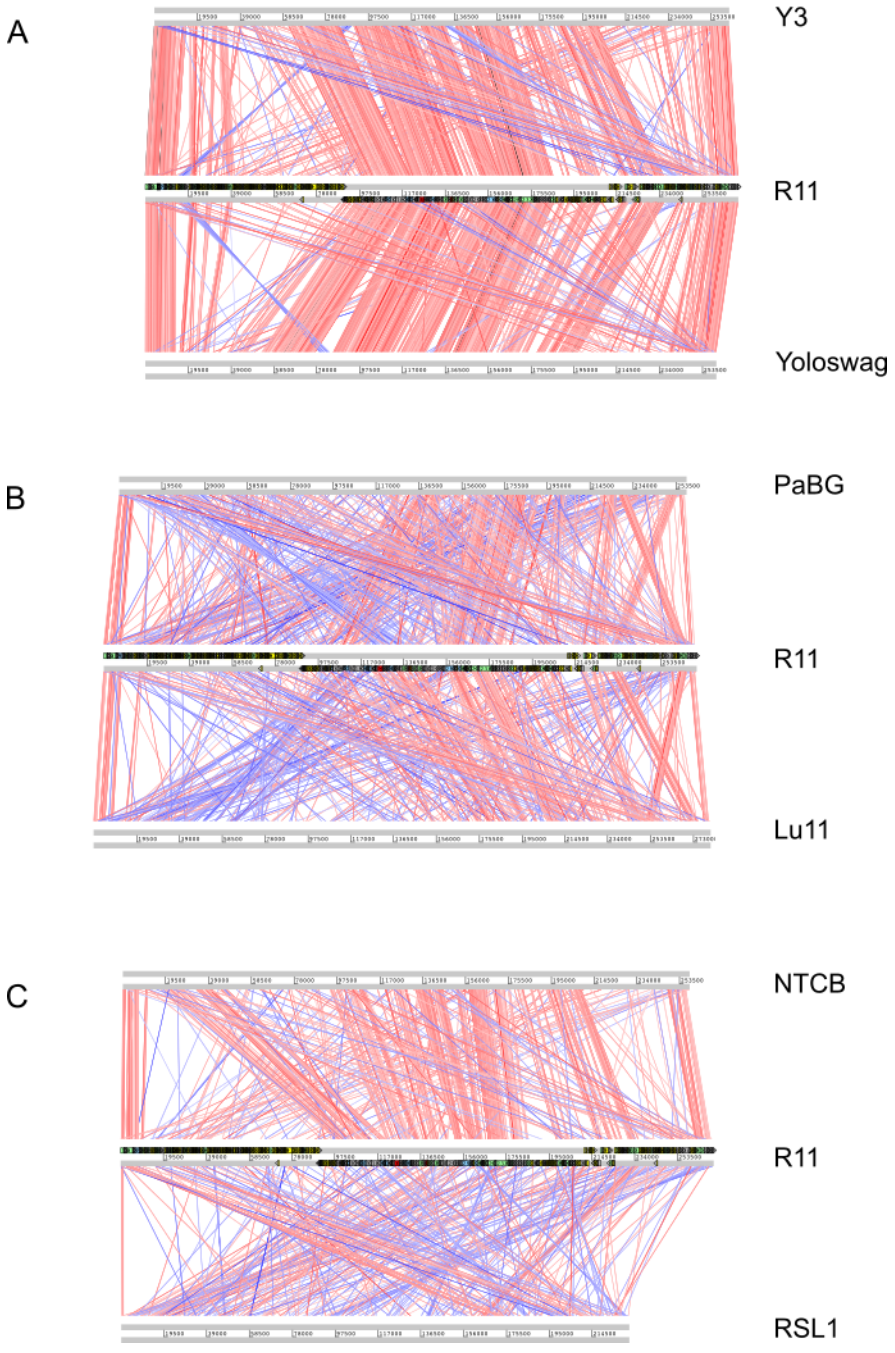



**Figure 3.16.** Genome comparison of ΦR11 and related phages suggest plastic genomes.

**A** Phages Y3 and yoloswag compared to ΦR11 which have the highest degree of amino acid homology in their products compared to other phages. **B** and **C** show *Pseudomonas* phages, PaBG and Lu11, phage NTCB and *Ralstonia* phage RSL1, respectively. Genomes of all phages are orientated to begin at the same homologous protein, a putative helicase. Images were generated using ACT, using a tblastx

comparison with a cut off value of $1x10^{-4}$. Red bar signifies areas of high amino acid sequence homology with blue bars showing high homology of inverted amino acid sequences.

### 3.2.6.2 TEM analysis of ΦR11-like phages

TEM images of phage Y3 and Lu11 show what has been described as hair-like structures on their tails (Buttimer et al., 2018; V. Shaburova a et al., 2006). Only one other group of sequenced phages has been reported to have these hair-like structures which are the Rak2-like phages that are jumbo phages and includes members such as *P. carotovorum* phage CBB (Buttimer et al., 2017b) and coliphage 121Q (NC_025447.1, (Hua, 2016)). TEM images of ΦR11 virions show that they may contain such hair-like appendages around their tails as well (**Figure 3.17**). Image resolution however, is not of a high enough quality to definitely conclude that these phages are "hairy". Given the high amino acid sequence identity ΦR11 shares with Y3 it is very likely that it is also a hairy jumbo phage. ΦR11 has a head size of roughly 120 nm with an extended tail of 180 nm.

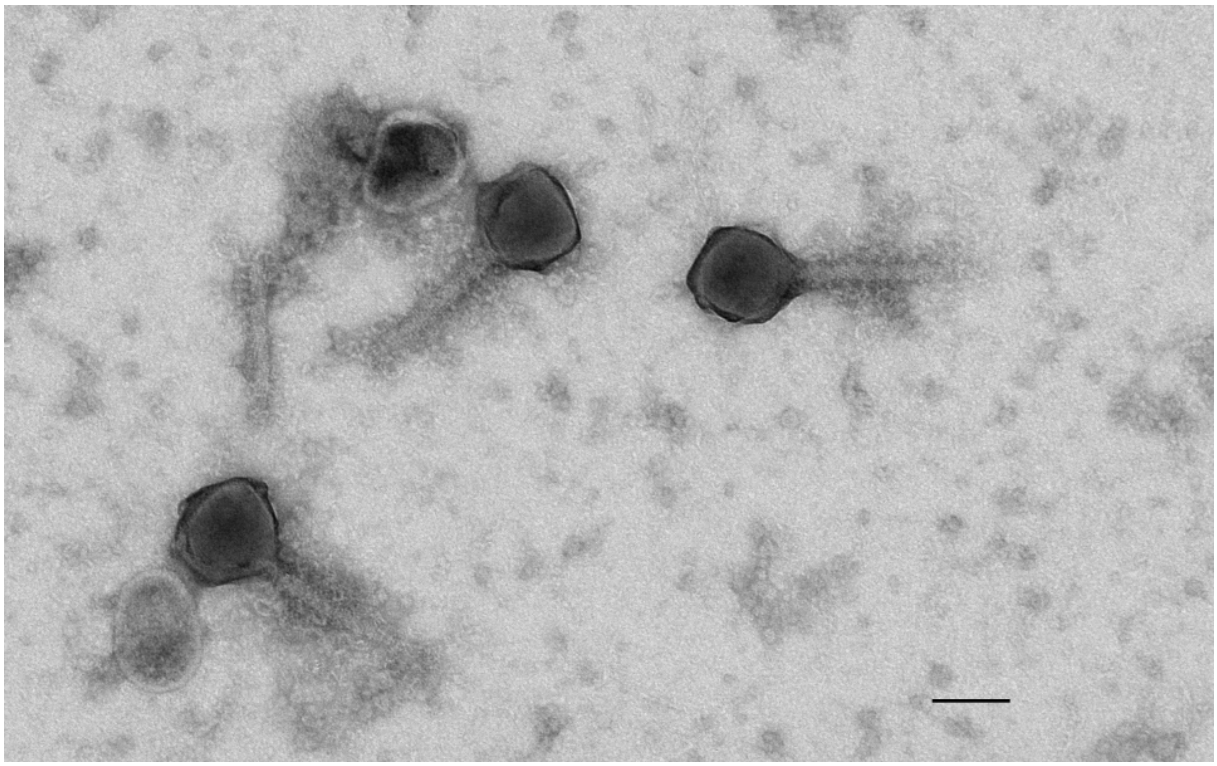


**Figure 3.17** TEM images of jumbo phage ΦR11.
Three intact ΦR11 virions are visible in the image. Image quality is not high enough to accurately determine if there are hairy appendages but there are thin fibre like strands visible from areas around the tail that may be hairs. The black bar represents 100 nm.

### 3.2.7 Predicted receptor and host ranges of the *P. atrosepticum* phages

To predict the putative receptor of the 36 sequenced *P. atrosepticum* phages, they were tested on two mutant strains of SCRI1043 (**Table 3.10**). These were a flagellar mutant, TER2, which was previously shown to be resistant to ΦTE, and SCC34 an uncharacterized LPS mutant, that is resistant to ΦM1. Phages that were unable to infect TER2 but are able to infect SCC34 were considered flagella-dependent. On the other hand, if a phage could not infect SCC34, but could infect TER2, then it was considered LPS dependent. Phages that could not infect either strain were either considered to recognise an alternative receptor, or because the LPS mutant, SCC34 is uncharacterized, another area of the LPS.

Results from the ΦTE-like phages show that they are flagella dependent since they did not infect the flagellar mutant but were able to infect the LPS mutant. This follows the same pattern as ΦTE which was previously shown to be flagella-dependent (Blower et al., 2012a). The ΦM1-like phages also follow the same pattern as ΦM1 i.e. they are able to infect the flagellar mutant, but not the LPS mutant. The ΦR1-like phages can infect both flagellar and LPS mutants. This suggests that their receptor is either an alternative exposed area of the LPS or another type of receptor, such as an outer membrane protein. The ΦRC6-like phages showed a mixed pattern where ΦRC6 cannot infect either the LPS or flagellar mutants. However, the other two ΦRC6-like members, ΦRC50 and ΦRC62 could not infect the LPS mutant but could infect the flagellar mutant. ΦRC6 has two predicted tail fibres which will need to be compared with those of ΦRC50 and ΦRC62 to determine if the difference in infectivity could be due to subtle differences in the amino acid sequence. It is possible that ΦRC6 recognises a different exposed area of the LPS, thus is capable of infecting SCC34. Both the jumbo phages are unable to infect TER2 but are able to infect SCC34, suggesting that they are both flagella dependent.

The 36 *P. atrosepticum* phages were also tested on a library of *P. atrosepticum* and *P. carotovorum* strains (**Table 3.10**). The ΦTE-like phages tended to be highly selective, with all capable of infecting only one other *P. atrosepticum* strain, (other than SCRI1043) which is SCRI1044. Interestingly ΦTE-like phages that belong to either subgroup one or two could infect an additional *P. atrosepticum* strain (SCRI41). The

ΦM1-like phages generally had a much broader host range with all members capable of also infecting SCRI1044 and SCRI1064. Furthermore, ΦM1-like phages belonging to subgroups one to four could infect three other *P. atrosepticum* strains, SCRI41, SCRI13 and SCRI1064. Members belonging to subgroup five could only infect SCRI1044 and SCRI1064. The ΦR1-like phages also had a broad host range, capable of infecting six *P. atrosepticum* strains. Curiously, a few of the ΦR1-like phages appeared to infect one or two of the *P. carotovorum* strains. The ΦRC6-like phages showed a more selective host range with all members capable of infecting only two other *P. atrosepticum* strains. ΦRC6 itself could infect an additional *P. atrosepticum* strain. Another interesting result was that all ΦRC6-like phages could infect a *P. carotovorum* strain. Jumbo phage ΦRC63 was also quite selective and was only capable of infecting one other *P. atrosepticum* strain. However, it was also shown to be able to infect two other *P. carotovorum* strains. ΦR11 had a slightly broader host range than ΦRC63 and is capable of infecting six *P. atrosepticum* and two *P. carotovorum* strains. Overall the *P. atrosepticum* phages can infect 11 different strains of pectobacteria.

**Table 3.10** Receptor and host range results of *P. atrosepticum* phages

| | | | *P. atrosepticum* | | | | | | | | | | | | | | *P. carotovorum* | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | Subgroup | Phages | SCC34 | TER2 | SCRI1043 | SCRI1 | SCRI24 | SCRI39 | SCRI41 | SCRI13 | SCRI1001 | SCRI1034 | SCRI1044 | SCRI1055 | SCRI1056 | SCRI1058 | SCRI1064 | SCRI101 | SCRI112 | SCRI114 | SCRI122 | SCRI132 | SCRI198 | SCRI103 | SCRI106 | SCRI115 | SCRI116 | SCRI124 | SCRI127 | SCRI130 | SCRI166 | SCRI192 |
| TE-like | 1 | S34 | Y | X | Y | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S61 | Y | X | Y | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | 2 | A2 | Y | X | Y | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | A5 | Y | X | Y | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S22 | Y | X | Y | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | 3 | A4S | Y | X | Y | X | X | X | X | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | A4L | Y | X | Y | X | X | X | X | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | A6 | Y | X | Y | X | X | X | X | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S21 | Y | X | Y | X | X | X | X | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S41 | Y | X | Y | X | X | X | X | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S72 | Y | X | Y | X | X | X | X | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | B1 | Y | X | Y | X | X | X | X | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| M1-like | 1 | B24 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | 2 | M5 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | 3 | A1 | X | Y | Y | X | X | X | Y | Y | x | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | M4 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | 4 | M2 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S32 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S33 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S42 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S64 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | S71 | X | Y | Y | X | X | X | Y | Y | X | Y | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | 5 | RC9 | X | Y | Y | X | X | X | X | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | RC10 | X | Y | Y | X | X | X | X | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| R1-like | | R1 | Y | Y | Y | X | X | X | Y | Y | X | Y | Y | Y | X | X | Y | X | X | X | X | X | Y | X | X | X | X | Y | X | X | X |
| | | RC58 | Y | Y | Y | X | X | X | Y | Y | X | Y | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | Y | X | X | X |
| | | RC64 | Y | Y | Y | X | X | X | Y | Y | X | Y | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | RC41 | Y | Y | Y | x | X | X | Y | Y | X | Y | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | Y | X | X | X | X |
| | | RC30 | Y | Y | Y | X | X | X | Y | Y | x | Y | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | RC36 | Y | Y | Y | X | X | X | Y | Y | X | Y | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | Y | X | X | X |
| | | RC33 | Y | Y | Y | X | X | X | Y | Y | X | Y | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | | RC18 | Y | Y | Y | X | X | X | Y | Y | X | Y | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| RC6-like | | RC6 | X | X | Y | X | X | X | X | X | X | X | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | Y | X | X | X |
| | | RC50 | X | Y | Y | X | X | X | X | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | Y | X | X | X |
| | | RC62 | X | Y | Y | X | X | X | X | X | X | X | Y | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | Y | X | X | X |
| RC63-like | | RC63 | Y | X | Y | X | X | X | X | X | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | Y | Y | X | X |
| R11-like | | R11 | Y | X | Y | Y | Y | X | X | X | X | Y | Y | Y | X | X | Y | X | X | X | X | X | X | X | X | X | X | X | Y | Y | X | X |

Y (green) signifies a positive infection with X (blue) a negative infection.

### 3.2.8 Discussion

Prior to this study there were only five fully sequenced genomes of *P. atrosepticum* phages and to address the paucity of genomic information on *P. atrosepticum* phages, 36 phages that were isolated on SCRI1043 were sequenced. Characterisation of this library of phages showed that they could be divided into six different groups. 24 of these phages were classed as ΦTE-like or ΦM1-like given their high similarity to the corresponding phages. ΦTE and ΦM1 were both shown to be transducing phages and are highly efficient at transferring chromosomal or plasmids markers in *P. atrosepticum* (Blower et al., 2012a; Toth et al., 1997). As the ΦTE-like and ΦM1-like phages are the most common group found and as they are similar to the two transducing phages, they should be screened for their ability to transduce. This observation may suggest that transducing phages of *P. atrosepticum* may be relatively common and this highlights the need to screen all phages for their ability to transduce, should they be used for future phage therapy cocktails. While the ΦTE-like and ΦM1-like phages are similar to each other there are many subtle difference in their genomes that may account for their different phenotypes. For instance, the host ranges of the ΦTE-like group phages are different within the group and similarly for the ΦM1-like group. Similar to the majority of sequenced phages, most predicted proteins from both groups of phages are hypothetical, with no known function.

The remaining 12 phages from the library were spilt into four different groups. One of these groups, the ΦR1-like phages, showed high similarity to a previously submitted *P. carotovorum* phage, phage Jarilo (MH059637, unpublished). Sequence comparisons show that even though Jarilo was isolated on a different bacterial host, *P. carotovorum* DSM30170, certain ΦR1-like viruses shared higher sequence homology to Jarilo compared to other members within the group. Host range results of the ΦR1-like phages also show that certain members could also infect either one or two *P. carotovorum* strains. One explanation for this is that, as the genomes of the *P. carotovorum* strains were never sequenced and the strains described by only biochemical assays, it could be that they were misclassified. Alternatively, they could be classified correctly and then this could mean that some ΦR1-like phages can indeed infect *P. carotovorum*. This could be a logical assumption to make given the isolation

of phage Jarilo, a very similar phage isolated on a *P. carotovorum* strain. Phages have been historically named after the host on which they have been isolated (Adriaenssens and Brister, 2017). This was done in the past because phages were originally considered to be highly specific for their host. This was a reason for their use to identify specific bacterial strains by phage typing (Clark and March, 2006). However, naming phages in this way has given the impression that phages isolated on their respective hosts are propagating on that host in the environment. The ΦR1-like phages are referred to as *P. atrosepticum* phages, while phage Jarilo would be referred as a *P. carotovorum* phage. However, theoretically if one of the susceptible *P. carotovorum* strains had been used to isolate ΦR1-like phages then they would also be classed as *P. carotovorum* phages. In reality, it is likely that the ΦR1-like phages are in fact replicating on different *P. atrosepticum* and *P. carotovorum* strains in the environment, and indeed other species of unknown bacteria.

The ΦRC6-like group is another of the new groups of *P. atrosepticum* phages sequenced. They were shown to belong to the *Ounavirinae* subfamily showing sequence homology to the founding phages of the four genera in the subfamily. One of the interesting features discovered when analysing their genomes was a potential CRISPR repeat. In bacteria, these repeats have DNA fragments (known as spacers) separating them, that are homologous to phage genomes and they function as an adaptive anti-phage system to recognise phages with the same DNA as phages from previous infections. However, it has also been shown that *V. cholerae* phages have been shown to carry their own CRISPR-Cas systems (Seed et al., 2013). These phages encode CRISPR-Cas systems that interfere with another antiphage system in their host, *V. cholerae* called phage-inducible chromosomal islands (PICI), that impedes phage replication (Naser et al., 2017). Analysis of the two spacers in ΦRC6 revealed no homologous sequences in NCBI although only about 7% of all known spacers have a sequence homology with published sequences in the databases, with the other 93% of spacers showing no sequence homology to anything (Shmakov et al., 2017). It is unknown whether the CRISPR repeats found in ΦRC6 are true repeats as there are no CRISPR-Cas homologs found in the phage. However most of the genomes encode hypothetical proteins with many of the proteins having no significant hits so it is possible that there are proteins that can interact with the predicted CRISPR repeats. *P. atrosepticum* strains have also been reported to carry a CRISPR-Cas

system so perhaps the repeats can be used instead by a host encoded system (Staals et al., 2016).

The final two phages, ΦRC63 and ΦR11 are two different jumbo phages which are both flagellum-dependent. ΦRC63 shows no nucleotide similarity to any sequenced phage, while ΦR11 shows a low degree of nucleotide homology to two previously isolated jumbo phages. However, both phages share many protein homologs with other jumbo phages previously submitted to the databases. As these jumbo phages have been sequenced very recently there is little information, apart from bioinformatic analyses or TEM images. There is information on the receptor for one of the jumbo phages, *Salmonella* phage SPN3US, which was found to be flagella-dependent. This concurs with the receptor data for ΦRC63 and ΦR11 that suggests that both the *P. atrosepticum* phages are dependent on flagella (Lee et al., 2011). Jumbo phages are currently thought to represent a relatively rare group of phages with very few genomes deposited in databases (Yuan and Gao, 2017). This makes the genomic data of ΦRC63 and ΦR11 particularly valuable in helping understand this rare group of large phages. It also suggests that we have perhaps only revealed the tip of an iceberg with the phage sequences analysed in this chapter and there must be considerable diversity of phage genomes yet to be discovered.

# Chapter 4 Type III Toxin antitoxin mediated abortive infection in *P. atrosepticum* phages

## 4.1 Introduction

The archetypal Type III toxin antitoxin system, $ToxIN_{Pa}$, was serendipitously discovered on a cryptic plasmid, pECA1039, originating from the phytopathogen *P. atrosepticum* SCRI1039 (Fineran et al., 2009). It has since been shown to be an effective antiphage system against several *P. atrosepticum* phages (Fineran et al., 2009). The first *P. atrosepticum* phage to be characterised in detail with regard to abortive infection, is the transducing flagella-dependent *Myoviridae* member, ΦTE. This phage was isolated from sewage effluent originating from the Milton sewage works in Cambridge on an uncharacterised LPS mutant, SCC34 of *P. atrosepticum* (Fineran et al., 2009). ΦTE is aborted by the $ToxIN_{Pa}$ system but able to evolve rare escape mutants that are no longer aborted. These escape phages were found to have an expansion of a repeat region in their genomes that was able to code for a functional "pseudo-antitoxin"(Blower et al., 2012a). As well as being the first *P. atrosepticum* phage studied with regards to Type III TA mediated Abi it was also the first *P. atrosepticum* phage that was genomically sequenced. The second *P. atrosepticum* phage to be studied with regards to Type III TA mediated Abi is ΦM1. This a transducing, LPS-dependent *Podoviridae* member that was isolated in 1988 from samples obtained at the Finham sewage works in Coventry (Toth et al., 1997). ΦM1 is aborted not just by $ToxIN_{Pa}$ but also by another Type III TA system, ($TenpIN_{Pl}$) which originates from the chromosome of *Ph. luminescens*, when cloned and expressed on a plasmid in *P. atrosepticum* (Blower et al., 2009, 2017). Escape mutants of ΦM1 from both TA systems were found to have mutations in a single ORF encoding a toxic protein of unknown function (Blower et al., 2017).

Escape mutants of both ΦTE and ΦM1 provided little information on the possible mechanism of Type III TA system activation in *P. atrosepticum* phages. This chapter attempts to address this matter and covers the impacts of *P. atrosepticum* harbouring Type III TA systems on multiple *P. atrosepticum* phages. This includes the phages described in the previous chapter plus further characterisation of ΦTE and ΦM1, the original aborted *P. atrosepticum* phages.

## 4.2 ΦTE-like phages require an antitoxin repeat to escape the ToxIN$_{Pa}$ system

ΦTE was the first phage characterised that could escape the ToxIN$_{Pa}$ system. All ΦTE escapes that were insensitive to the ToxIN$_{Pa}$ system had an expanded region that could code for an RNA antitoxin mimic. Thus far there have been no other escape mutations identified so it is still not clear if there are any other means of escape. The ΦTE-like phages described in the previous chapter had been previously screened for sensitivity to the ToxIN$_{Pa}$ system by Tim Blower during his PhD study but at that time there was no knowledge of the genomic nature of the phages. He found that some phages that are now known to be ΦTE-like were all aborted by the ToxIN$_{Pa}$ system. However, none of these phages tested were able to evolve escapes. The ΦTE-like phages were tested again in this study with titres above $10^{10}$ pfu.ml$^{-1}$ and this confirmed that the previous data were accurate (data not shown). However, as these phages were genomically similar to ΦTE this result was both curious and unexpected. The genome region that contains the source of antitoxin repeats in ΦTE escape mutants was compared to the annotated representative members from each of the subgroups, ΦS34, ΦA2 and ΦA4s (**Figure 4.1**). Closer analysis of the region revealed that in all three of the subgroups, the potential antitoxin mimic region was partially or completely deleted.

ΦTE contains a short region in its genome that matches closely with approximately 1.5 of the ToxI$_{Pa}$ repeat. However, this region is out of phase with ToxI$_{Pa}$ so that even when transcribed will not code for the complete functional antitoxin repeat necessary to neutralise ToxI$_{Pa}$ (**Figure 4.2A**). In order to escape ΦTE has to produce a completely functional ToxI$_{Pa}$ RNA subunit. It is thought that ΦTE is able to evolve escape mutants at a very low frequency ($1x10^{-8}$), due to strand slippage during replication that causes the 1.5 times ToxI$_{Pa}$ mimic region to be expanded to either 4.5 times or 5.5 times (Blower et al., 2012a). These phage mutants which can produce a complete "pseudo ToxI$_{Pa}$" mimic enable resistance to abortive infection by ToxIN$_{Pa}$, presumably by neutralising ToxN$_{Pa}$. Phages ΦS34 and ΦA2, are not able to escape as they containing only 0.5 times of a potential antitoxin mimic and thus, even if strand slippage occurs, they can never produce a complete functional ToxI$_{Pa}$ mimic (**Figure 4.2**). For phage ΦA4s this region is completely deleted and there is no genomic sequence that closely resembles ToxI$_{Pa}$ (**Figure 4.1**).

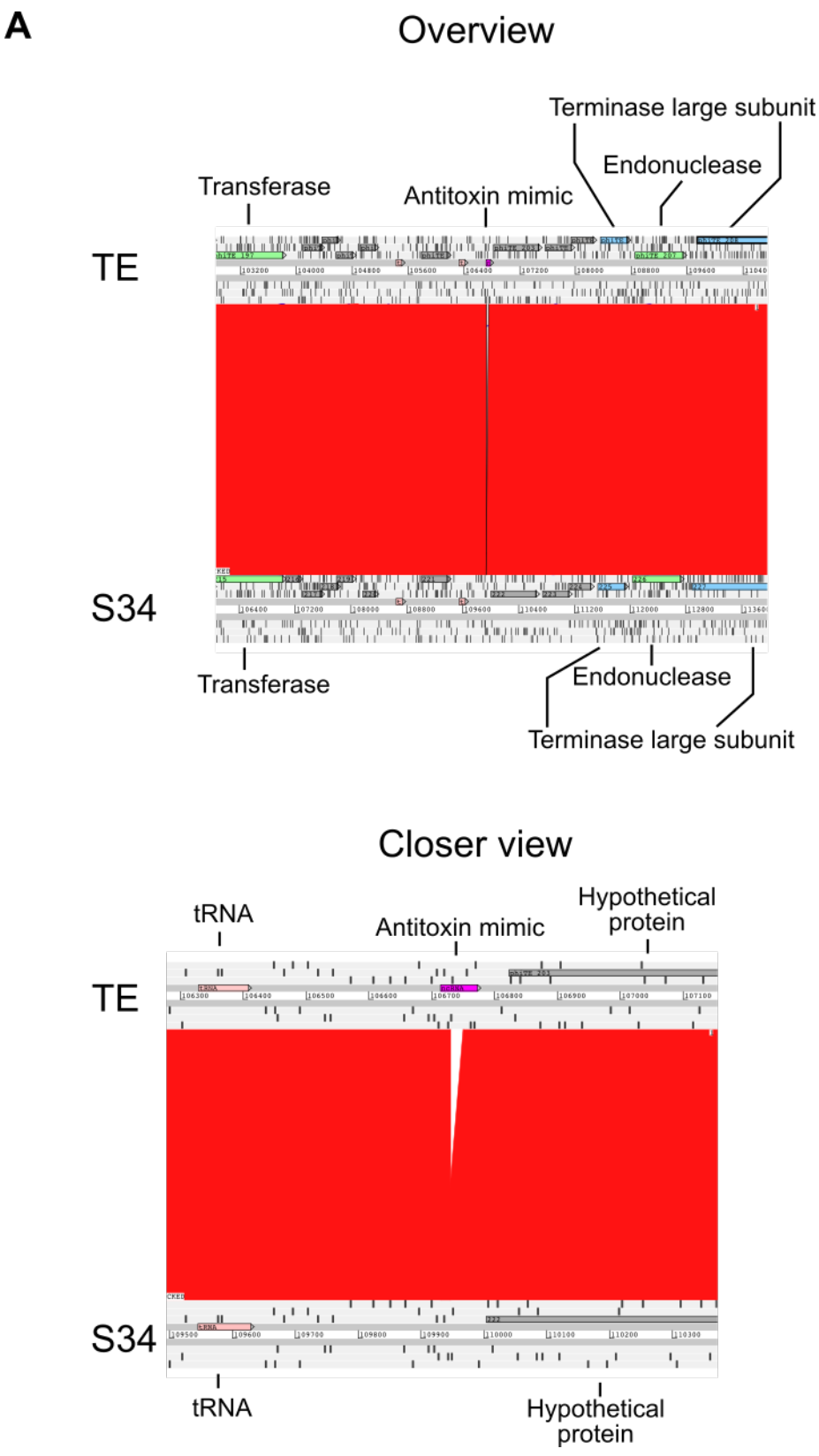



**Figure 4.1** continued overleaf

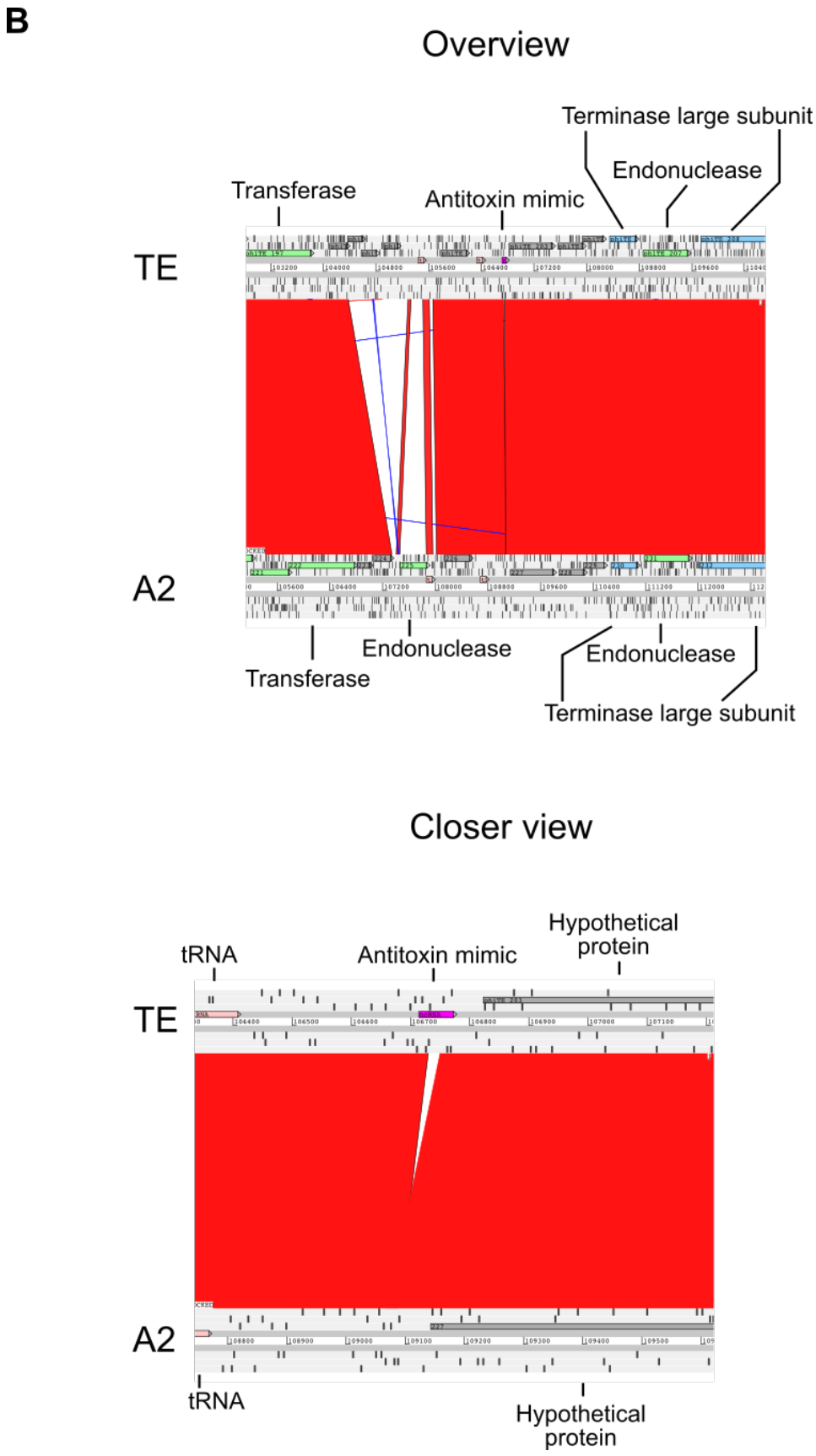



**Figure 4.1** continued overleaf

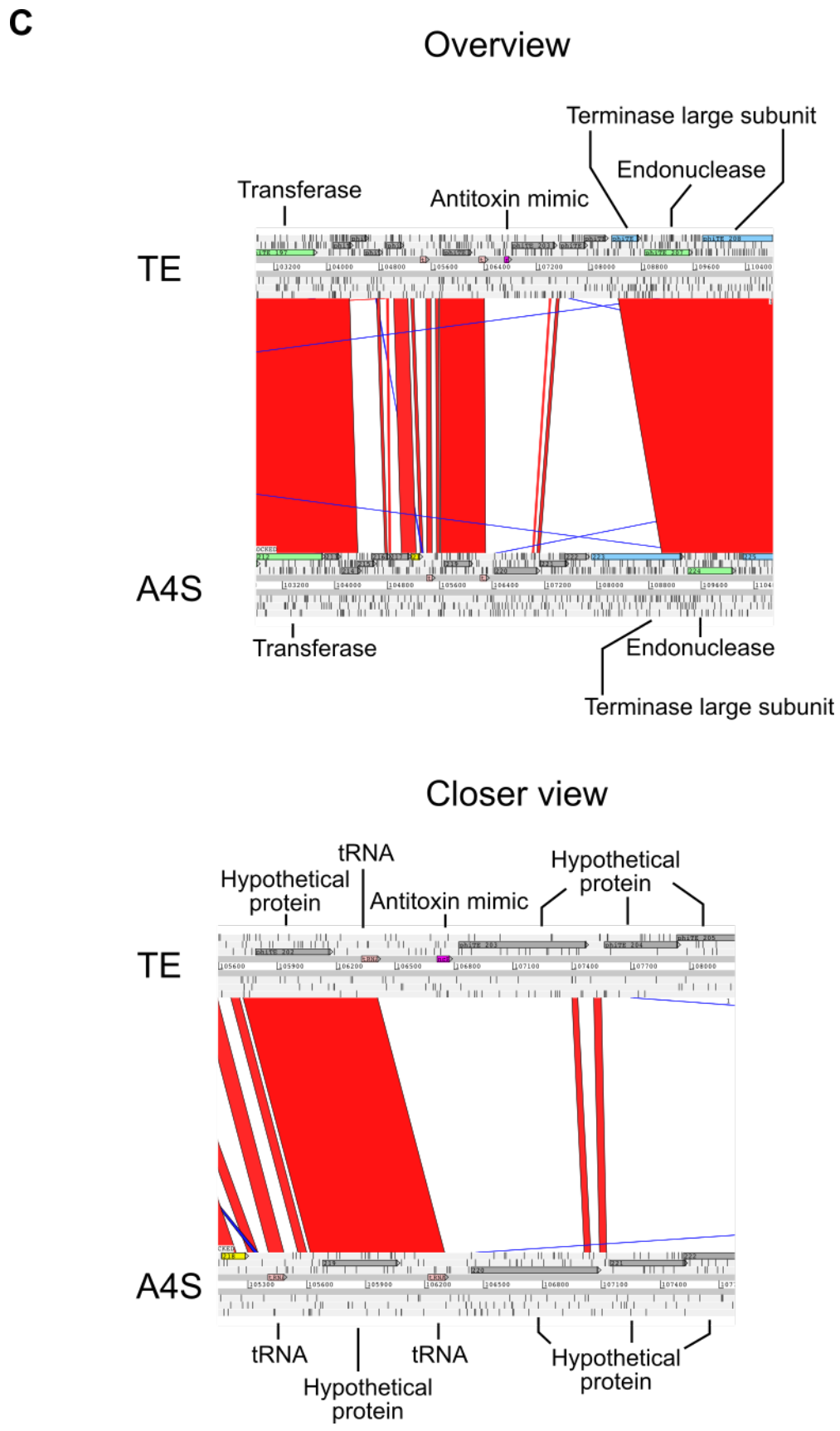



**Figure 4.1** Figure legend overleaf

**Figure 4.1** ΦTE escape locus compared to the same region of phages, ΦS34, ΦA2 and ΦA4s.

**A**. ΦTE's escape locus compared to ΦS34. Both phages share high nucleotide similarity in this region, however there is an insertion of a small nucleotide sequence in ΦTE compared to ΦS34 and ΦA2 that allows it to code for a 1.5 times antitoxin mimic. **B**. ΦTE's escape locus compared to ΦA2. ΦA2 shares less homology in this region with ΦTE encoding ORFs with products of unknown functions (grey boxes) that are not present in ΦA2. ΦA2 also contains an endonuclease that is not present in ΦTE. The antitoxin mimic region for ΦA2 is also the same as in ΦS34 with ΦTE containing a small nucleotide insertion. **C**. ΦTE's escape locus compared to ΦA4s which show the least nucleotide homology for the three subgroup representatives. ΦA4s does not have any nucleotide sequence that shares any homology with the antitoxin mimic sequence found in ΦTE. Comparison files were generated using a blastn comparison with a $1x10^{-4}$ cut off value. Red bars show nucleotide sequences of high homology with blue bars representing inverted nucleotide sequences of high homology.

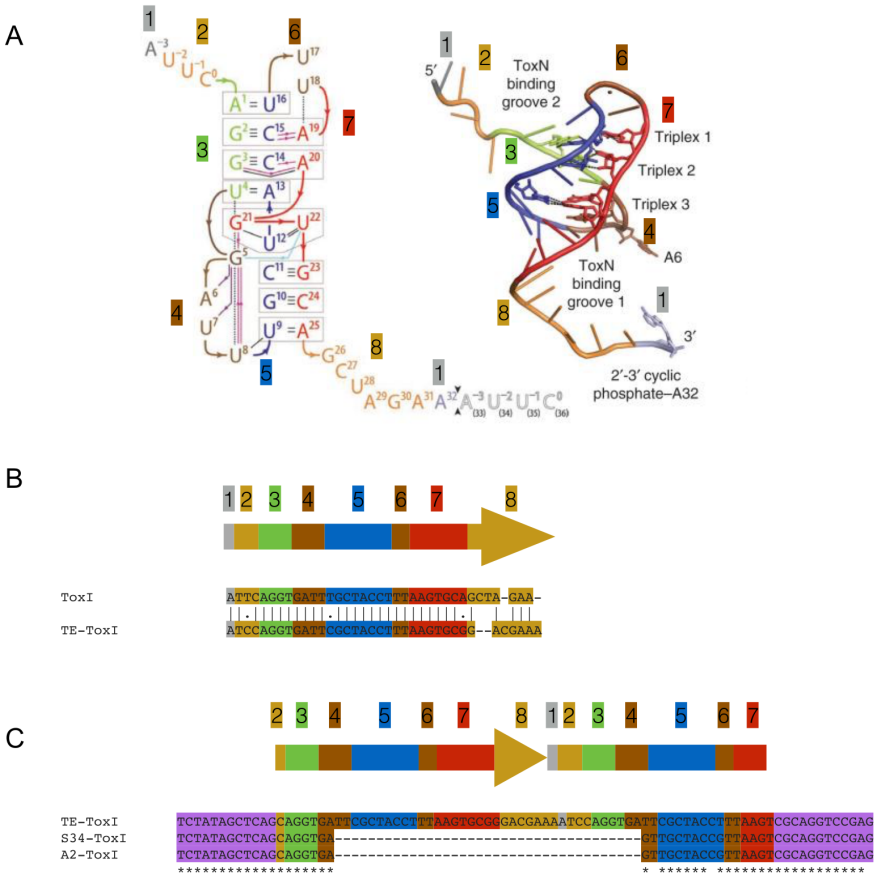



**Figure 4.2** Phages ΦS34 and ΦA2 do not contain a complete ToxI$_{Pa}$ mimic.

**A**. The ToxI$_{Pa}$ RNA pseudoknot subunit structure required to neutralise ToxN$_{Pa}$. Taken from (Blower et al., 2011). **B** The 36 nucleotide long ToxI$_{Pa}$ sequence required to neutralise ToxN$_{Pa}$ aligned with the ToxI$_{Pa}$ mimic produced by ΦTE escapes. **C**. The nucleotide sequence of the ToxI$_{Pa}$ mimic region of ΦTE, ΦS34 and ΦA2. Wild type ΦTE contains a 1.5 times ToxI$_{Pa}$ mimic repeat that is out of phase with the ToxI$_{Pa}$. When transcribed the region produces either a ToxI$_{Pa}$ mimic with regions 2-8 or 1-7 and thus a completely functional ToxI$_{Pa}$ subunit is not produced. Expansion of this region however, allows a complete ToxI$_{Pa}$ mimic to be produced. ΦS34 and ΦA2 do not contain a complete repeat and when transcribed contains the regions 2-7. As they do not contain regions 1 or 8, these phages do not have the potential to encode a complete functional unit. The colours and numbers are coordinated in **A**, **B** and **C** and represent the related areas in each case.

## 4.3 Abortive infection in ΦM1-like phages

### 4.3.1 Extensive analysis of ΦM1 escapes show mutations in a single toxic ORF

ΦM1 was the second characterised phage that could escape the ToxIN$_{Pa}$ system. The original ΦM1 escapes sequenced had mutations in a single ORF, M1-23, encoding a toxic protein of unknown function (Blower et al., 2017). As all 10 mutations were unique it was though that other mutations in the M1-23 protein must also be able to allow escape. To characterise this protein further, 51 newly isolated escapes were sequenced, 35 of which were by a rotation student, Samuel Kidman. Collectively, all 51 newly isolated escapes, the 10 original escapes, and other escapes sequenced throughout this ΦM1 phage study are shown here for completeness (**Table 4.1**). This led to the discovery of 21 different mutations (**Figure 4.3**). The escape locus of M1 is also shown in **Figure 4.4** with all the different mutations mapped onto the Figure which is drawn to scale. With the addition of the new mutations, mutations in all three bases in the start codon have been observed, confirming that the start site of M1-23 had been correctly assigned. Interesting escapes of ΦM1 include ΦM1-E11 which has a mutation that causes an extremely N-terminal truncation resulting in a two amino acid long peptide. This, and other N-terminal truncations, show that the M1-23 protein is not essential for the replication of ΦM1. The majority of ΦM1 escapes contained missense mutations leading to substitutions, mostly in the C-terminal domain amino acids. This implies that the C-terminal region is important for the ability of the protein to activate ToxIN$_{Pa}$. Furthermore, escapes ΦM1-E48 and ΦM1-E49 (both the same mutation) have a mutation in the stop codon leading to a 10 amino acid extension. The addition of this short amino acid sequence must somehow interfere with the ability of M1-23 to activate ToxIN$_{Pa}$, presumably again by disrupting the integrity of the C-terminal region of the protein.

ΦM1 was additionally shown to be aborted by another Type III TA system, TenpIN$_{Pl}$, when expressed in SCRI1043. Again, as in the ToxIN$_{Pa}$ system, escapes from TenpIN$_{Pl}$ had the same mutations in the M1-23 protein. This was possibly a surprising result, as ΦTE (which is more strongly aborted by ToxIN$_{Pa}$) is completely insensitive to TenpIN$_{Pl}$ (Blower et al., 2017). However, the results for ΦM1 suggests that activation of both the ToxIN$_{Pa}$ and TenpIN$_{Pl}$ system can also be linked. All ΦM1 experiments performed here were published during the period of this PhD (Blower et al., 2017).

**Table 4.1** M1-23 mutations

| Phages | Mutation relative to ΦM1 wt | Location relative to *phiM1-23* | Effect on M1-23 | Effect on M1-22 | Number of phages |
|---|---|---|---|---|---|
| ΦM1-E1<br>ΦM1-E2<br>ΦM1-E3<br>ΦM1-E4 | 15169 A to C | A1C | Met to Leu | Tyr to Ser | 4 |
| ΦM1-C<br>ΦM1-E5<br>ΦM1-E6<br>ΦM1-E7<br>ΦM1-E8<br>ΦM1-E9<br>ΦM1-U4<br>ΦM1-U5<br>ΦM1-U6<br>ΦM1-U9 | 15170 T to C | T2C | Met to Thr | Tyr to Tyr | 10 |
| ΦM1-E10 | 15171 G to A | G3A | Met to Ile | Asp to Asn | 1 |
| ΦM1-E11 | 15175 A to T | A7T | Lys to Stop | Gln to Leu | 1 |
| ΦM1-E12 | 15256 A to T | A88T | Lys to Stop | Glu to Val | 1 |
| ΦM1-X<br>ΦM1-E13<br>ΦM1-E14<br>ΦM1-E15 | 15288 AA to A | AA120A | FS to Stop after 9aa | FS to Stop after 45aa | 4 |
| ΦM1-B<br>ΦM1-E16<br>ΦM1-E17<br>ΦM1-E18<br>ΦM1-E19<br>ΦM1-E20 | 15292 C to T | C124T | Arg to Stop | Pro to Leu | 6 |
| ΦM1-E21 | 15359 T to C | T191C | Leu to Pro | - | 1 |
| ΦM1-E22<br>ΦM1-E23 | 15384 T to G | T216G | Tyr to Stop | - | 2 |
| ΦM1-Y | 15397 G to A | G229A | Asp to Asn | - | 1 |
| ΦM1-W | 15398 A to T | A230T | Asp to Val | - | 1 |
| ΦM1-E24<br>ΦM1-E25 | 15398 A to C | A230C | Asp to Ala | - | 2 |
| ΦM1-U7 | 15398 A to G | A230G | Asp to Gly | - | 1 |
| ΦM1-O<br>ΦM1-Q<br>ΦM1-E26<br>ΦM1-E27 | 15407 A to C | A239C | Gln to Pro | - | 4 |
| ΦM1-D<br>ΦM1-E | 15410 T to C | T242C | Met to Thr | - | 2 |
| ΦM1-E28 | 15410 T to G | T242G | Met to Arg | - | 1 |
| ΦM1-V<br>ΦM1-E29 | 15415 T to G | T247G | Tyr to Asp | - | 2 |
| ΦM1-A<br>ΦM1-E30 | 15416 A to C | A248C | Tyr to Ser | - | 2 |

**Table 4.1** continued. M1-23 mutations

| Phages | Mutation relative to ΦM1 wt | Location relative to *phiM1-23* | Effect on M1-23 | Effect on M1-22 | Number of phages |
|---|---|---|---|---|---|
| ΦM1-Z<br>ΦM1-E31<br>ΦM1-E32<br>ΦM1-E33<br>ΦM1-E34<br>ΦM1-E35<br>ΦM1-E36<br>ΦM1-E37<br>ΦM1-E38<br>ΦM1-E39<br>ΦM1-E40<br>ΦM1-E41<br>ΦM1-E42<br>ΦM1-E43<br>ΦM1-E44<br>ΦM1-E45<br>ΦM1-E46<br>ΦM1-U1<br>ΦM1-U2<br>ΦM1-U8<br>ΦM1-U10 | 15416 A to G | A248G | Tyr to Cys | - | 21 |
| ΦM1-E47 | 15419 T to C | T251C | Leu to Pro | - | 1 |
| ΦM1-E48<br>ΦM1-E49 | 15424 T to C | T256C | Stop to Arg (extension of 10 aa) | - | 2 |
| Different mutations: 21 | | | Total mutations: 70 | | |

FS – frame shift
aa – Amino acid
Stop – stop codon

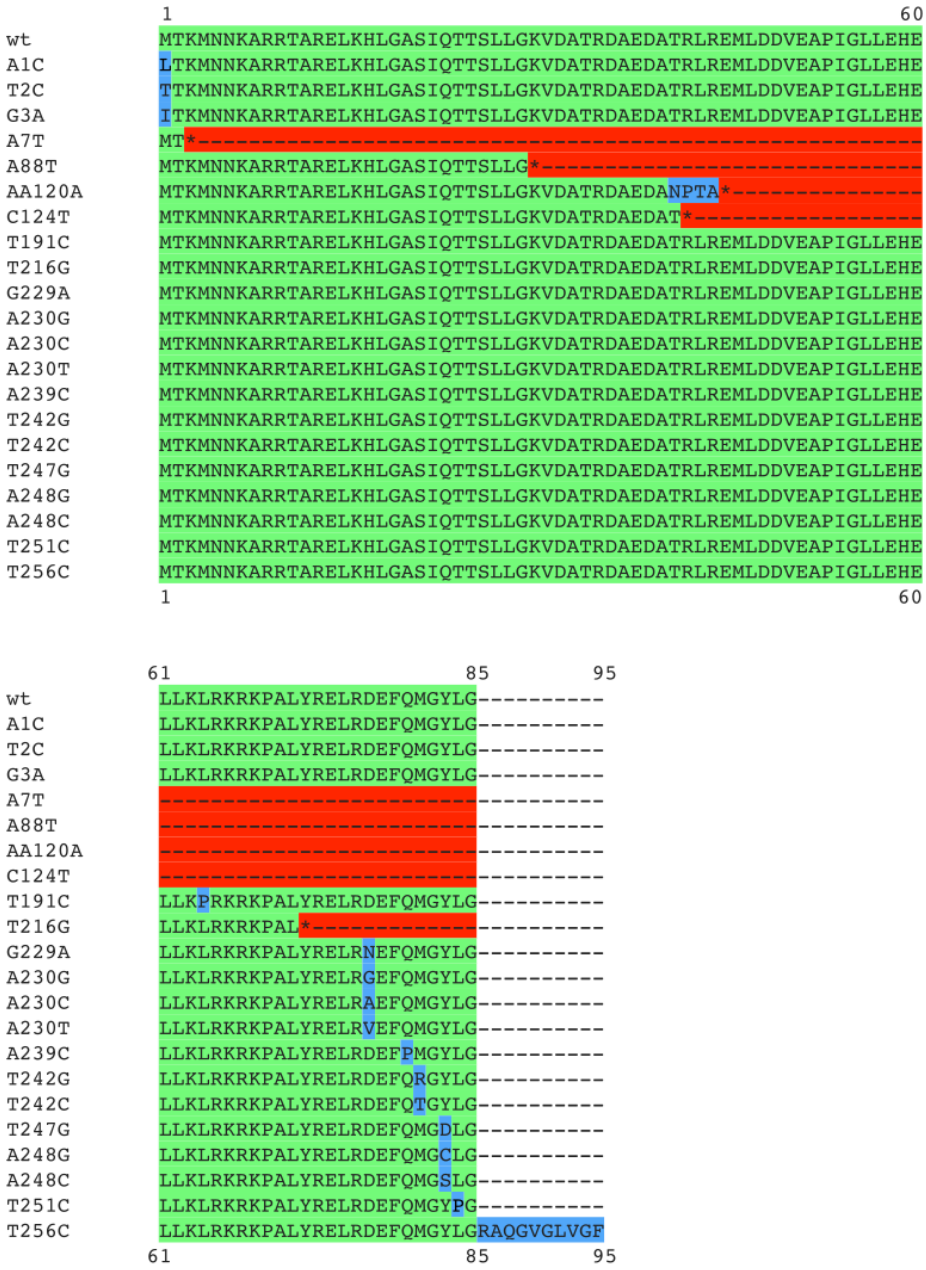



**Figure 4.3** Alignments of M1-23 and escape variants.

Amino acids that are identical to the wild type protein are highlighted in green. Amino acids that are highlighted in blue represent differences from the wild type protein.

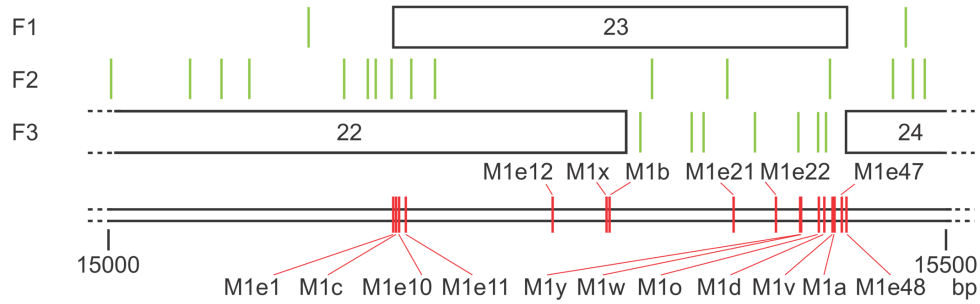



**Figure 4.4** The escape locus of ΦM1.

Taken from (Blower et al., 2017). The figure is drawn to scale. Green bars represent stop codons. Red bars represent the locations of all identified mutations and a representative phage escape mutant harbouring that mutation. As *phiM1-22* overlaps with *phiM1-23*, certain N-terminal mutations also affect the *phiM1-22* product (**Table 4.1**). The phiM1-23 product has an unknown function while *phiM1-22* and *phiM1-24* encodes a predicted DNA exonuclease and endonuclease, respectively.

### 4.3.2 ΦRC10 is highly sensitive to TenpIN$_{Pl}$ and escapes via a M1-23 homologue

ΦRC10 is an M1-like phage that was found to be highly sensitive to TenpIN$_{Pl}$ but, in terms of EOP, insensitive to ToxIN$_{Pa}$ (**Table 4.2**). The genome of ΦRC10 contains a M1-23 homologue called RC10-26 which differs by only a single amino acid (**Figure 4.5**). Previously described phages that showed high homology to ΦM1 were analysed and their M1-23 homologs aligned (**Figure 4.5**). The M1-23 homologs are largely conserved including the amino acid residues found to be important in ToxIN$_{Pa}$ activation of ΦM1. Only one phage, PPSW1, has differing amino acids in residues important for ToxIN$_{Pa}$ activation. As the sequence is fairly conserved it may mean that M1-23 homologues confer some advantage in the environment as they are not necessary for replication in lab conditions.

In order to assess whether RC10-26 was involved in TenpIN$_{Pl}$ activation, 16 spontaneous escapes were isolated and their RC10-26 sequences determined (**Table 4.3**). The data showed they contained similar mutations to those found in ΦM1 ToxIN$_{Pa}$ escapes. There were nine different mutations with six mutations already seen in ΦM1 escape. Three of the mutations however, were novel, and this may suggest that, even though a large number of ΦM1 escapes had already been sequenced, it may be possible to generate more escape mutations.

ΦRC10 did not activate ToxIN$_{Pa}$ (as determined by EOPs) but plaques picked from plates of ΦRC10 propagated on SCRI1043 containing ToxIN$_{Pa}$ contained mutations in RC10-26 (data not shown). Therefore, it is unclear why ΦRC10 is more sensitive to TenpIN$_{Pl}$ than ΦM1, or why it is not aborted by ToxIN$_{Pa}$ even though ΦRC10 phages propagated on ToxIN$_{Pa}$ have mutations in RC10-26.

**Table 4.2.** EOP data of ΦM1 and ΦRC10

| Phage | TA system | EOP |
|---|---|---|
| ΦM1 | $ToxIN_{Pa}$ | $1.3x10^{-5}$ |
| | $TenpIN_{Pl}$ | $1.1x10^{-2}$ |
| ΦRC10 | $ToxIN_{Pa}$ | 0.64 |
| | $TenpIN_{Pl}$ | $1.3x10^{-5}$ |

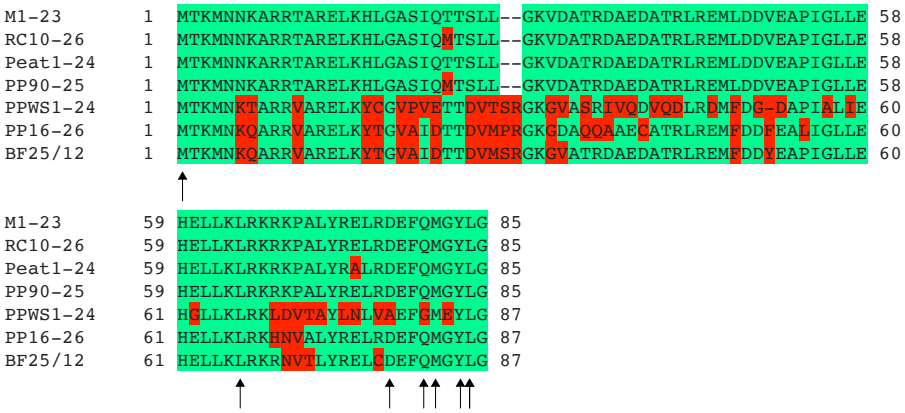



**Figure 4.5** M1-23 alignment with homologous proteins.

Green highlighted amino acid represents conserved residues. Red highlighted bases represent differences. Arrows show residues that are substituted in escape variants of M1-23.

**Table 4.3** ΦRC10 escape mutations

| Phage | Location relative to *rc10-26* | Effect on RC10-26 | Effect on RC10-25* | Number of phages | Found in ΦM1 escapes |
|---|---|---|---|---|---|
| ΦRC10a | A1T | Met to Leu | Tyr to Phe | 1 | No |
| ΦRC10b | T2C | Met to Thr | Tyr to Tyr | 1 | Yes |
| ΦRC10g | T161C and G244A | Ile to Thr and Gly to Arg | - | 1 | No |
| ΦRC10e | T242G | Met to Arg | - | 1 | Yes |
| ΦRC10h ΦRC10k | G242GG | Tyr to Val (extends protein by 66 aa) | - | 2 | No |
| ΦRC10m ΦRC10n | T247G | Tyr to Asp | - | 2 | Yes |
| ΦRC10d | A248C | Tyr to Ser | - | 1 | Yes |
| ΦRC10i ΦRC10j ΦRC10l ΦRC10o ΦRC10p | A248G | Tyr to Cys | - | 5 | Yes |
| ΦRC10c ΦRC10f | T256C | Stop to Arg (extension of 10 aa) | - | 2 | Yes |
| Different mutations: 9 | | | Total mutations: 16 | | |

*Similar to ΦM1, *rc10-26* overlaps with *rc10-25* which encodes a putative DNA exonuclease.

Aa – amino acid

Stop – stop codon

### 4.4 Type III TA mediated Abi infection in other *P. atrosepticum* phages

The other groups of *P. atrosepticum* phages were screened against both the $ToxIN_{Pa}$ or the $TenpIN_{Pl}$ system. Phage titres of around $10^{10}$ pfu.ml-1 were used in all the screens. Results from the screen showed that the ΦR1-like group (ΦRC63 and ΦR11) were not aborted by SCRI1043 carrying either $ToxIN_{Pa}$ or $TenpIN_{Pl}$. In contrast, ΦRC6-like group were strongly aborted by the $ToxIN_{Pa}$ system but remained unaffected by the $TenpIN_{Pl}$ system. Attempts were made to isolate $ToxIN_{Pa}$ escapes from the ΦRC6-like group but no putative escape plaques could be propagated.

### 4.5 Discussion

This chapter looks at the interaction of Type III TA systems with the six different groups of *P. atrosepticum* phages described in the previous chapter. Three of the six groups were aborted by one or more Type III TA systems. The other three remaining groups remained insensitive to Type III TA mediated abortive infection. The three groups that were sensitive to Type III TA systems were the ΦTE-like, ΦM1-like and ΦRC6-like group. None of these phages have any genes in common so it is unclear why these phages activate Type III TA systems while the other insensitive phages do not. Even within the three sensitive groups, Type III TA system activation seems to differ. The ΦTE-like and ΦRC6-like groups are highly sensitive to abortive infection from the $ToxIN_{Pa}$ system but are insensitive to the $TenpIN_{Pl}$ system, suggesting that different Type III TA systems have different mechanisms of activation. However, ΦM1 is aborted by both systems and is able to escape both through mutations in a single protein, M1-23, suggesting that the mode of activation of the Type III TA system may also be linked, in this phage. Another interesting observation is that ΦRC10 (a ΦM1-like phage) shows high sensitivity to the $TenpIN_{Pl}$ system but doesn't seem to be aborted by $ToxIN_{Pa}$. The M1-23 homolog in ΦRC10 only differs by a single amino acid which may account for the different sensitivity. ΦRC10 escapes of either the $TenpIN_{Pl}$ or $ToxIN_{Pa}$ system have mutations in the M1-23 homolog (RC10-26) that map to similar mutations previously identified in ΦM1 escapes. Thus it seems that this protein plays an important role in Type III TA system activation. Interestingly however, homologues of this protein are not found in ΦTE-like to ΦRC6-like phages confirming that there must be other routes to Type III mediated Abi system.

Data from the ΦTE-like phages suggest that activation of $ToxIN_{Pa}$ at least for ΦTE-like phages occurs through an essential phage process. The ΦTE-like phages are unable to evolve to escape the $ToxIN_{Pa}$ system and only ΦTE itself (which may represent a natural rare variant) has the potential to generate an antitoxin mimic. Even then, escape mutants of ΦTE occurs at an extremely low frequency ($1x10^{-8}$) with the majority of the wild type ΦTE population being aborted (Blower et al., 2012a). ΦRC6-like phages also exhibit a similar phenotype in that none are able to evolve escapes. This suggests that activation of $ToxIN_{Pa}$ may target essential genes or that the very rare mutations required for escape are not selectable at the titres used. It is possible that several mutations are required, however very high titres are used for the assays (around $10^{10}$) so this is unlikely. Yet as powerful as $ToxIN_{Pa}$ seems in this case these phages are completely insensitive to the $TenpIN_{Pl}$ system.

The results of the ΦM1 and ΦRC10 experiments tell a completely different story from that of the ΦTE-like and ΦRC6-like phages. In ΦM1 and ΦRC10 phages both the $ToxIN_{Pa}$ and $TenpIN_{Pl}$ systems abort the phages, albeit to different degrees and yet escapes for both phages are through mutation of a single non-essential protein. This is in stark contrast to the ΦTE-like and ΦRC6-like data suggesting that Type III TA system activation is specific and may even target an essential process. A large number of ΦM1 and ΦRC10 escapes have been isolated and in all cases, the same protein (M1-23 or RC10-26) was mutant suggesting that this protein plays a key role in activating two Type III TA systems. Homologues of this protein are also found in other closely related phages that infect other bacterial species, including *P. carotovorum* and *Dickeya* (**Figure 4.5**). Many of the homologues are conserved suggesting that the protein is functional in these phages so it is likely that this protein plays an important role in viral replication under certain environmental conditions. The M1-23 protein was also shown to be toxic, even in the absence of $ToxIN_{Pa}$, and when expressed in a different host *i.e. E. coli* (Blower et al., 2017). As M1-23 is toxic, this suggests that it must be interacting with the different bacteria although the nature of this interaction is still a mystery.

The ΦR1 group phages are not aborted by either the $ToxIN_{Pa}$ or $TenpIN_{Pl}$ system, although they are in the same subfamily as the ΦM1-like phages. They do not contain a M1-23 homologue. However, neither does the ΦTE-like nor ΦRC6-like phages. The

two jumbo phages, ΦRC63 and ΦR11 also do not activate the $ToxIN_{Pa}$ or $TenpIN_{Pl}$ systems. There is very little information on jumbo phages but it is thought that the advantage that these phages have through carrying the large number of genes is that they don't have to rely as heavily on host encoded proteins to replicate. This may result in less interaction with the host which may be the reason why they are not aborted.

The work in this chapter leads to the following observations. Firstly, responses to Type III TA systems are different among the different groups of phages, even within the same group (ΦM1 and ΦRC10) where the Type III systems can be activated to differing extents. Furthermore, the activation mechanism for Type III TA systems may differ between systems yet can also be linked (TE is aborted by the $ToxIN_{Pa}$ but not $TenpIN_{Pl}$ system while ΦM1 is aborted by both systems and escapes through mutations in a single gene). Finally, there is no gene or common protein product common between phages that activate the Type III TA system. Different proteins may be interacting with the Type III TA system directly or different proteins interfere with host factors that can cause Type III TA activation. The M1-23 protein is toxic even in the absence of $ToxIN_{Pa}$ suggesting that its normal function involves interaction with the host rather than a direct interaction with $ToxIN_{Pa}$.

# Chapter 5 ToxIN$_{Pa}$ abortive infection in coliphage ΦCHAI8

## 5.1 Introduction

The work in the previous chapter showed that the Type III TA systems are able to cause abortive infection in phages from widely different genetic backgrounds. It also showed that a Type III TA system from another bacterium (*Ph. luminescens*) is able to abort phages in *P. atrosepticum*. Previous work has showed that ToxIN$_{Pa}$, when cloned and expressed on a plasmid in *E. coli*, is also able to abort range of coliphages (Blower et al., 2012b). Unfortunately, while there have been numerous attempts in this laboratory to isolate coliphage escapes from the ToxIN$_{Pa}$ and TenpIN$_{Pl}$ systems, none of the phages isolated that were aborted could escape (unpublished).

Although no coliphage escapes of ToxIN$_{Pa}$ and TenpIN$_{Pl}$ were isolated prior to this study, the Moineau laboratory that works on another Type III TA system, AbiQ, managed to isolate a coliphage that is aborted by AbiQ but additionally able to evolve escapes (Samson et al., 2013). The AbiQ system derives from *Lactococcus lactis* and is able to abort many lactococcal phages but, like ToxIN$_{Pa}$ and TenpIN$_{Pl}$ could still function as an antiphage system when transferred into a different host, such as *E. coli.* One of the coliphages (phage 2, a *V5virus*) was aborted by *E. coli* carrying the AbiQ system but could evolve escapes. When analysed these escapes had mutations in a single ORF that encodes a putative DNA polymerase.

During the course of this study, nine coliphages were isolated from the river Cam (as part of an outreach demonstration to teach school children how to isolate phages from the environment). This constituted part of a programme called Authentic Biology which aims to encourage schools to collaborate with their local university on a biological project. The collection of coliphages isolated during the teaching of the demonstration were tested for their sensitivity to both the ToxIN$_{Pa}$ and TenpIN$_{Pl}$ system. One of these phages, ΦCHAI8, was aborted by the ToxIN$_{Pa}$ system but proved able to evolve, escapes as described in this chapter.

### 5.2 ΦCHAI8 is aborted by ToxIN$_{Pa}$ and is able to evolve escapes

The coliphages were isolated on DH5α from water samples that originated from either the Milton sewage works or the river Cam in 2016. These phages were named with the prefix ΦCHAI followed by a number. They were screened for sensitivity to both ToxIN$_{Pa}$ and TenpIN$_{Pl}$ by spotting a serial dilution of phages on a lawn of DH5α containing a plasmid with either pTA46 (ToxIN$_{Pa}$), pFR2 (TenpIN$_{Pl}$) or their non-functioning frame shift equivalents (**Figure 5.1**). Results from the screen identified one phage (ΦCHAI8) that was aborted by ToxIN$_{Pa}$ but could evolve escapes that were no longer sensitive to the ToxIN$_{Pa}$ system (**Table 5.1**). Isolated plaques were picked, purified and confirmed to breed true as escape mutants based on EOP assays. These escape mutants replicated as efficiently on the plasmid containing the functioning ToxIN$_{Pa}$ system as on *E. coli* carrying the non-functioning frame shift ToxIN$_{Pa}$ system (**Table 5.1**).

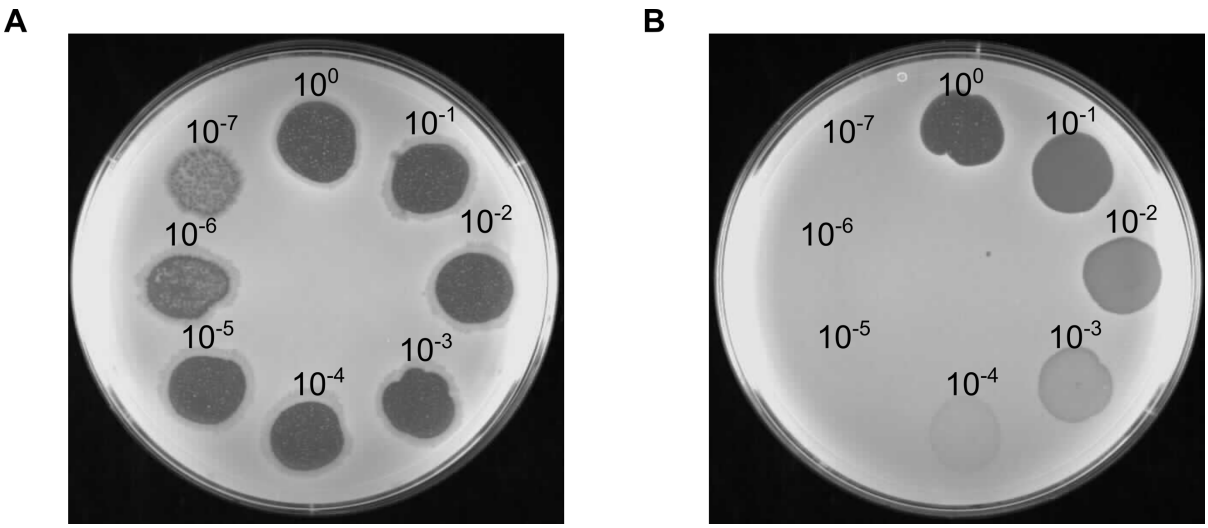



**Figure 5.1** Spot test showing ΦCHAI8 sensitivity to the ToxIN$_{Pa}$ system

Spot test of wild type ΦCHAI8 on a lawn of DH5α containing a plasmid with either **A**. a non-functioning ToxIN$_{Pa}$ system, pTA47 or **B**. a functioning ToxIN$_{Pa}$ system, pTA46. 10 µl spots of serially diluted phages were used.

**Table 5.1** EOPs of ΦCHAI8 and escape mutants

| **Phage** | **EOP vs ToxIN$_{Pa}$** | **EOP vs ToxIN$_{Pa}$ FS** |
|---|---|---|
| ΦCHAI8 | $1.4x10^{-6}$ | 1 |
| ΦCHAI8a | 0.028 | 1 |
| ΦCHAI8b | 0.39 | 1 |
| ΦCHAI8c | 0.42 | 1 |
| ΦCHAI8d | 0.06 | 1 |
| ΦCHAI8e | 0.26 | 1 |

### 5.3 ΦCHAI8 is a *V5virus*

Transmission electron microscopy images of ΦCHAI8 showed that it has a long contractile tail and thus belongs to the *Myoviridae* family (**Figure 5.2**). The genomic DNA of ΦCHAI8 was extracted and sequenced. The sequence revealed that ΦCHAI8 had a circular double stranded DNA genome of 138,889 base pairs (**Table 5.2**). Annotation of the genome predicted 211 ORFs and 7 tRNAs. A genome map of ΦCHAI8 is shown in **Figure 5.3**. Blastn results of the ΦCHAI8 genome showed that it belongs to the *V5virus* genus. This genus of phages was named after coliphage V5 which had little nucleotide sequence homology to any sequenced phages at that time (Santos et al., 2011). The V5 genome has not been published and all genome analysis of V5 is based on a descendent of the phage called rV5 (or recovered V5). This phage, rV5, formed part of a phage cocktail which was used to treat cattle infected with *E. coli* O157:H7. After the phage therapy treatment, phages (including rV) were recovered from the animals. As rV5 may have acquired mutations during its passage through the cattle it is considered a different phage (Kropinski et al., 2013). Nucleotide comparisons of ΦCHAI8 with other *V5viruses* showed that they have similar lengths, ORFs, tRNAs and nucleotide sequence identities with each other (**Table 5.2**). ACT generated images of ΦCHAI8 vs rV5 showed only very small insertions/deletions between the two phages (**Figure 5.4**). ΦCHAI8 also shows very high sequence homology with the genome of phage 2, which, as mentioned previously, is aborted by the AbiQ system but also able to evolve escapes (**Figure 5.4**).

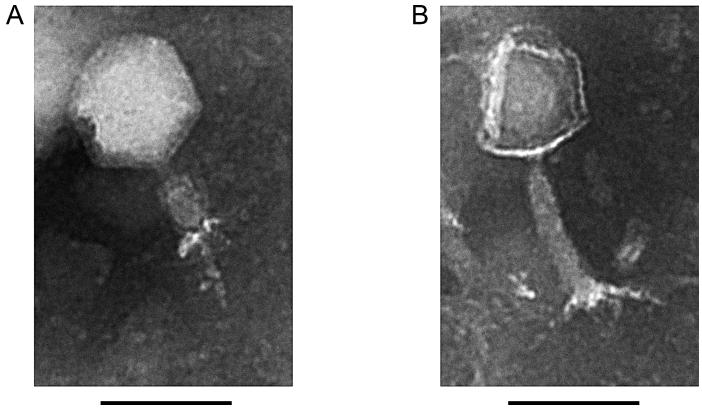



**Figure 5.2** Transmission electron microscope images of ΦCHAI8

TEM images of **A**. ΦCHAI8 with a contracted tail and **B**. with an extended tail. Black bars represent 100 nm.

**Table 5.2** Genomic analysis of ΦCHAI8 and comparisons to other *V5virus* members

| Phage | Genome size (bp) | ORFs | tRNAs | Identity to ΦCHAI8 (%) | Accension no. |
|---|---|---|---|---|---|
| ΦCHAI8 | 138, 889 | 215 | 7 | 100 | Unpublished |
| rV5 | 137, 947 | 233 | 6 | 93.1 | DQ832317.1 |
| *E. coli* 0157 typing phage 5* | 137, 973 | 214 | 6 | 92.3 | KP869103.1 |
| APCEc02* | 135, 400 | 219 | 5 | 89.7 | KR698074.1 |
| Phage 2 JES-2013 | 136, 910 | 220 | 4 | 91.3 | KC690136.1 |
| PDX* | 138, 427 | N/A | N/A | 91.4 | MG963916.1 |
| Slur12* | 136, 898 | 207 | 7 | 92.8 | LN881735.1 |
| Slur16* | 136, 896 | 206 | 7 | 92.8 | LN881727.1 |
| FV3 | 136, 947 | 218 | 5 | 87.9 | JQ031132.1 |
| Murica | 135, 391 | 212 | 7 | 92.2 | KT001917.1 |
| FFH2 | 139, 020 | 216 | 6 | 89.7 | KJ190158.1 |
| V18* | 129, 454 | 207 | 3 | 88.6 | KY683736.1 |

*Phage sequences that had to be reoriented to begin at the same location as ΦCHAI8 and rV5

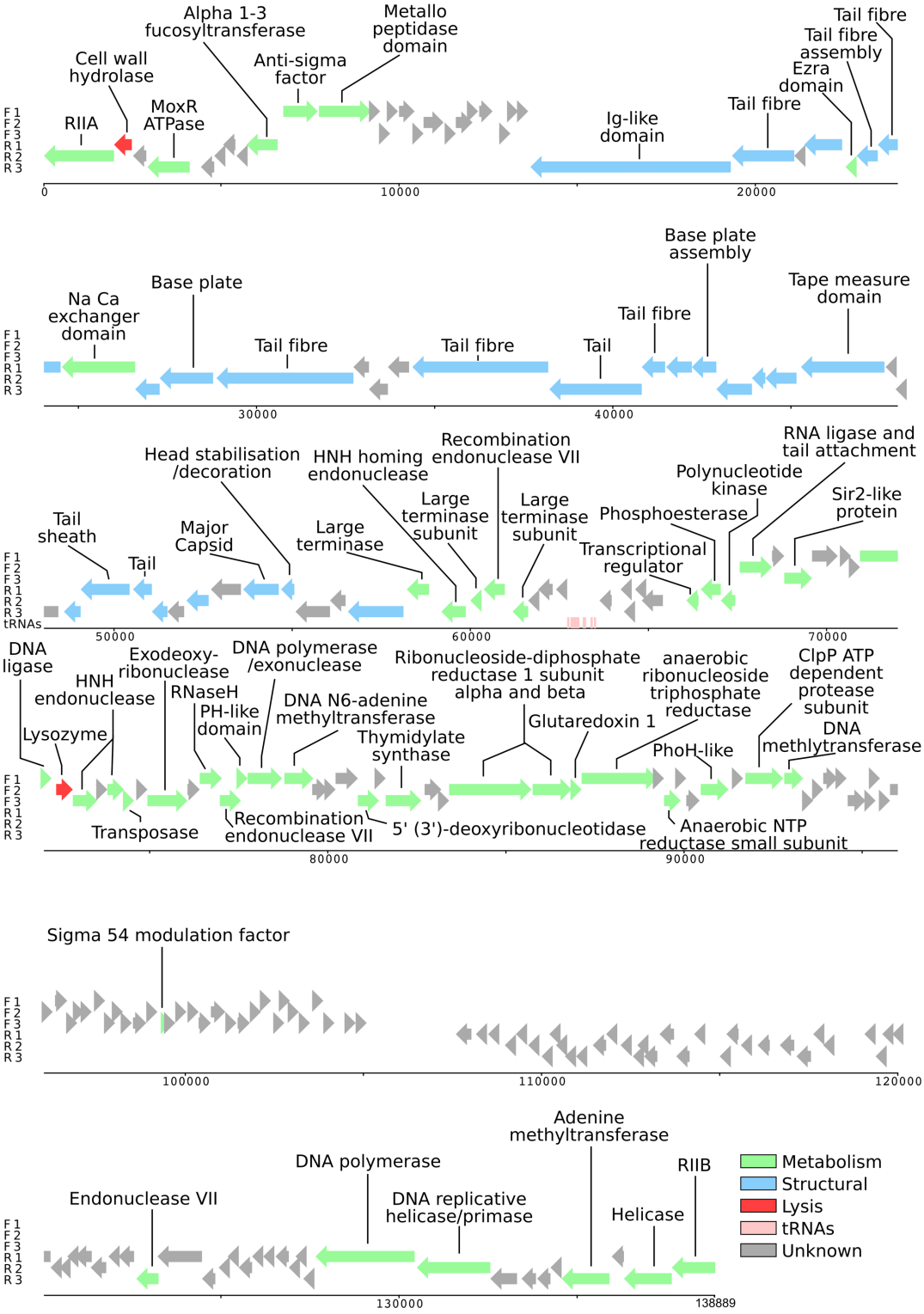



**Figure 5.3** Genomic map of ΦCHAI8

The genome has been orientated to begin at RIIA and with RIIB in accordance to other phages encoding RIIA and RIIB homologs. F1-3 show ORFs on the forward reading

strands split into their three different reading frames. R1-3 is the same but for the reverse strand. Arrows in green, blue and red, represent ORFs encoding proteins involved in metabolism, virion structure and lysis. Unlabelled structural proteins are homologs of PVP-SE1 structural proteins identified during a structural proteomics experiment. Grey arrows represent ORFs of unknown function with the location of tRNAs shown as pink boxes.

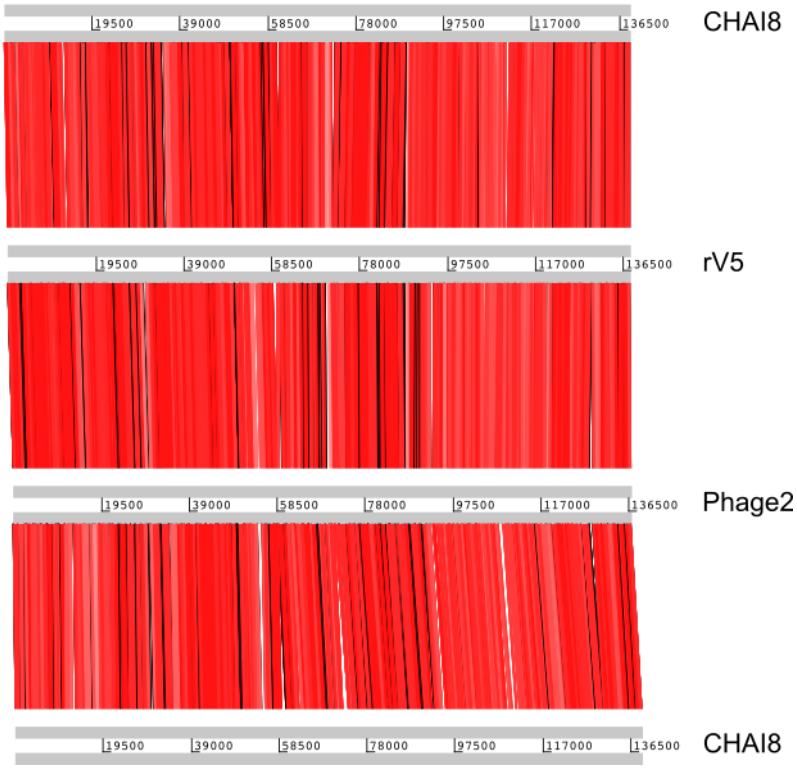



**Figure 5.4** ΦCHAI8 shows high similarity to both rV5 and Phage 2

Comparisons were generated using Artemis comparison tool using a tblastx comparison with a cut off value of $1x10^{-4}$. Red bars signify areas of high amino acid similarity.

## 5.4 ToxIN$_{Pa}$ escapes of ΦCHAI8 have deletions in a putative RNA escape locus

To determine the escape loci of ΦCHAI8, five independently isolated ToxIN$_{Pa}$ escapes, ΦCHAI8a-e, had their genomes sequenced. This revealed that all escapes had deletions of varying sizes in a common locus (**Figure 5.5**). The escape with the largest deletion (ΦCHAI8c) deleted 4,175 bases, disrupting seven proteins (**Table 5.3**). On the other hand, the escape containing the smallest deletion (ΦCHAI8d) showed only a single base deleted within an ORF, CHAI8-77. This ORF, CHAI8-77 was completely deleted in all other escapes, suggesting that it plays an important role in escaping the ToxIN$_{Pa}$ system. Homologs of CHAI8-77 found in other *V5viruses* have not been assigned a function, apart from containing a predicted AAA-ATPase domain. To predict the function of CHAI8-77, it was entered into a Phyre2 pipeline using the default settings. The Phyre2 program artificially models a structure using the submitted amino acid sequence and using the model, maps it onto similar known protein structures (Kelley et al., 2015). For CHAI8-77, the top hit mapped to the kinase domain of the T4 Pnk protein (**Figure 5.6**).

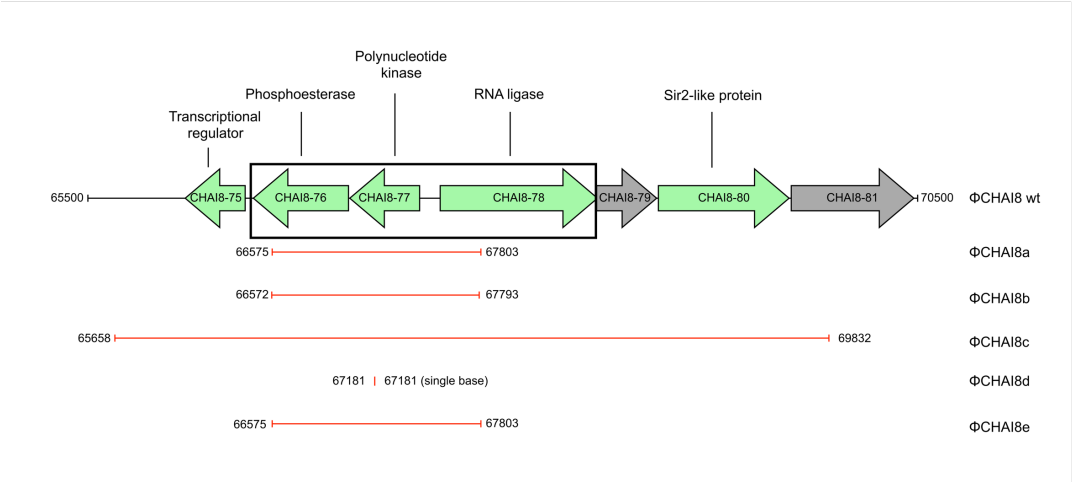



**Figure 5.5** The escape locus of ΦCHAI8 ToxIN$_{Pa}$ escapes

Red bars represent the nucleotide sequences deleted in escapes. The black box represents ORFs encoding proteins involved in the putative RNA repair system.

**Table 5.3.** ORFs within the ΦCHAI8 escape locus

| ORF | Start position | End position | Length (bp) | Size (aa) | E value | Significant BLASTP hit |
|---|---|---|---|---|---|---|
| CHAI8-75 | 66423 | 66073 | 351 | 116 | 1.00E-79 | putative transcriptional regulator [*Escherichia* phage 2 JES-2013] |
| CHAI8-76 | 67025 | 66471 | 555 | 184 | 4.00E-136 | putative phosphoesterase [*Escherichia* phage 2 JES-2013] |
| CHAI8-77 | 67440 | 67033 | 408 | 135 | 3.00E-91 | hypothetical protein [*Escherichia coli*] contains AAA and polynucleotide kinase domains |
| CHAI8-78 | 67559 | 68476 | 918 | 305 | 0.00E+00 | putative RNA ligase and tail attachment protein [*Escherichia* phage vB_EcoM_FFH2] [*Escherichia* phage slur16] |
| CHAI8-79 | 68473 | 68823 | 351 | 116 | 5.00e-79 | hypothetical protein Ec2_0077 [*Escherichia* phage JES2013] |
| CHAI8-80 | 68832 | 69593 | 762 | 253 | 0.00E+00 | putative Sir2-like protein [Enterobacteria phage vB_EcoM-FV3] [*Escherichia* phage slur16] |
| CHAI8-81 | 69607 | 70320 | 714 | 237 | 5.00E-173 | hypothetical protein [*Escherichia coli*] |

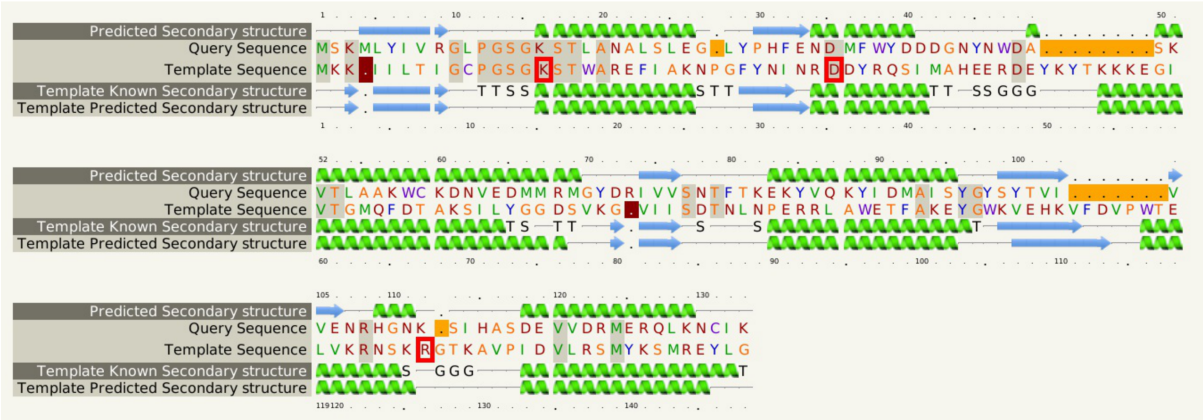



**Figure 5.6** Phyre2 alignment of CHAI8-77 has a similar fold to T4 Pnk

Phyre2 alignment of CHAI8-77 (query sequence) and the top hit, the kinase domain of T4 Pnk (template sequence). Conserved residues are highlighted in grey, deletions in ΦCHAI8 in orange and deletions in Pnk in red. Catalytic residues determined by the Catalytic Site Atlas are in red boxes (Furnham et al., 2014).

The Pnk protein (short for polynucleotide kinase) forms part of an RNA repair pathway in T4. It is a bifunctional protein capable of both kinase and phosphatase activities (Sirotkin et al., 1978). As part of an RNA repair pathway it is required to repair the 3'$PO_4$/5'OH ends of damaged RNA. The kinase activity cleaves the terminal 3'-$PO_4$ converting it to a 3'OH while the phosphate drives the phosphorylation of the 5'-OH end to 5'$PO_4$. To join the break, a second protein, Rnl1, a RNA ligase is able to ligate the "healed" gap (Amitsur et al., 1987). The Pnk-Rnl1 proteins are an evolutionarily conserved RNA repair pathway with homologs found in other viruses, bacteria, protozoans, yeasts and plants (Englert and Beier, 2005; Wang and Shuman, 2005; Wang et al., 2002).

Pnk is encoded by *pseT* and is a 301 amino acid long protein arranged in two domains. The N-terminal region of the protein specifically residues 1-147 contain the kinase domain with the C-terminal region i.e. residues 149-301 covering the phosphatase domain (Wang et al., 2002). CHAI8-77 is only 135 amino acids long and shows structural homology with the kinase domain of Pnk (**Figure 5.6**). A conserved amino acid sequence at position 9 to 16, GxxGxGKS, is present in the kinase domain of Pnk. This sequence is defined as a Walker A box motif which is sometimes referred to as a P loop. This motif (in particular the lysine residue) is essential for ATP binding (Hanson and Whiteheart, 2005). Two of the motifs in the Walker A box, Lys15 and Ser16, additionally form part of the active site. This active site is comprised of three other residues, Asp35, Arg38 and Arg126, involving a total of five residues (Wang and Shuman, 2002). CHAI8-77 may also contain a Walker A box but only three of the active site residues are conserved (**Figure 5.6**). There is a tryptophan and lysine residue in place of Arg38 and Arg126, respectively. Lysine, alike arginine is a positively charged residue, however this conservative substitution at both Arg38 and Arg126 still abolishes kinase activity in the T4 Pnk (Wang and Shuman, 2002). Conversely, in another related bifunctional kinase phosphatase protein, CPT, involved in chloramphenicol resistance in *Streptomyces venezuelae*, the Arg38 residue is replaced by an isoleucine suggesting that these residues need not be strictly conserved between the related proteins (Izard, 2001).

Both the kinase and phosphatase activities of Pnk are essential for its RNA repair function. As CHAI8-77 only matches the kinase domain, it is missing the fundamental

phosphatase domain. Interestingly, the upstream adjacent protein to CHAI8-77, CHAI8-76 is a predicted phosphoesterase which could provide the phosphatase activity and compensate for the lack of such in CHAI8-77. Further analysis revealed that CHAI8-76 is a member of the phosphoprotein phosphatase family, a type of serine/threonine phosphase that contains five conserved motifs (Barford et al., 1998; Lohse et al., 1995) (**Figure 5.7**). These five motifs play an important role in catalysis and substrate binding (Goldberg et al., 1995; Griffith et al., 1995; Zhuo et al., 1994). There has been a sixth motif proposed (Connelly and Leach, 2002) but generally five motifs are defined (Krogh et al., 2005; Williams et al., 2008). These phosphatases are part of a large evolutionarily conserved double stranded DNA repair pathway involving two proteins called the MR (Mre11-Rad50) complex. The phosphatase, Mre11 acts as a manganese-dependent nuclease with Rad50 associated with ATPase activity (Paull and Deshpande, 2014). Homologs of Mre11 and Rad50 are not only present in phages but also bacteria, archaea and eukaryotes (Connelly and Leach, 2002; Herdendorf et al., 2011; Paull and Deshpande, 2014; Reiter et al., 2002).

As mentioned previously the T4 RNA repair pathway requires both the Pnk protein and a RNA ligase, Rnl1 (Amitsur et al., 1987). ΦCHAI8 contains a Rnl1 homolog, located downstream and adjacent to CHAI8-77 called CHAI8-78 that is also commonly deleted in escapes. The Rnl1 proteins share six conserved motifs that make up a nucleotide-binding pocket that is also present in CHAI8-78 (**Figure 5.8**). Thus, together, CHAI8-76, CHAI8-77 and CHAI8-78 may be able to provide the necessary activities required to function as an RNA repair pathway.

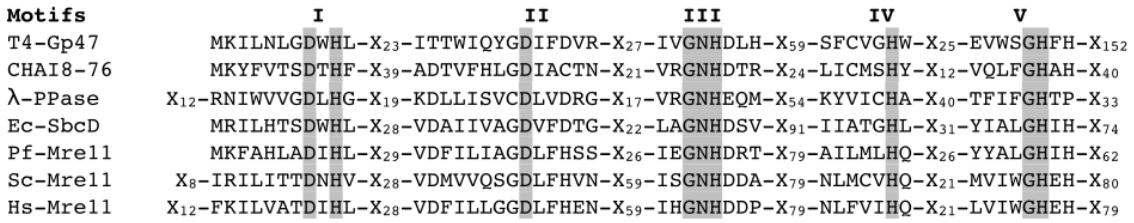



**Figure 5.7** CHAI8-76 shares five conserved motifs found in Mre11 proteins.

Conserved residues from all proteins are shaded in grey. The proteins that are included in the figure are two phages, T4 and λ, a bacterium, Ec (*E. coli*), an archaeon, Pf (*Pyrococcus furiosus*), a yeast, Sc (*Saccharomyces cerevisiae*) and human, Hs (*Homo sapiens*).

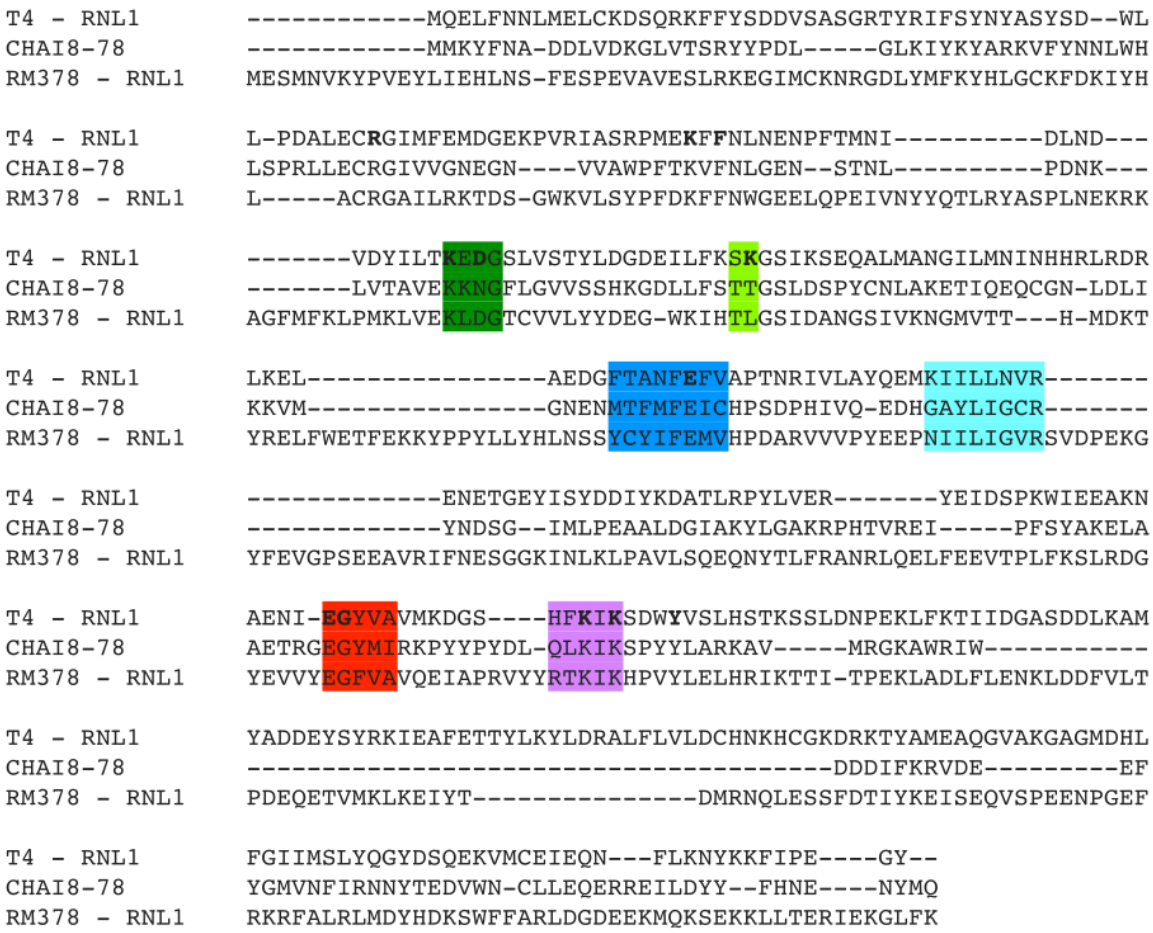



**Figure 5.8** Alignments of T4 Rnl1 with CHAI8-78 and another homologue from phage RM378.

The six nucleotidyl transferase motifs are shaded in dark green (I), light green (Ia), dark blue (III), light blue (IIIa), red (IV) and purple (V). In bold are residues that were shown to be essential for protein function in T4.

## 5.5 Discussion

This chapter focuses on one particular phage (ΦCHAI8) and mutants that escape the ToxIN$_{Pa}$ system. Other coliphages that had been previously isolated were either insensitive to ToxIN$_{Pa}$ or were aborted but couldn't escape and so ΦCHAI8 provided an invaluable new insight into ToxIN$_{Pa}$ activation in coliphages. Escapes of ΦCHAI8 insensitive to the ToxIN$_{Pa}$ system were shown to have deletions in genes encoding a putative RNA repair pathway. This is a novel escape locus that has not been associated with Type III TA system activation before. However, it is not the first time the Pnk and Rnl1 proteins have been implicated in abortive infection.

In T4, Pnk and Rnl1 prevents abortive infection by the Prr system. The Prr system forms a component of a two part antiphage defence in *E. coli*. The other component is the type Ic DNA restriction endonuclease EcoprrI. T4 is able to prevent restriction by EcoprrI through producing an inhibitor protein called Stp (Penner et al., 1995). However, *E. coli* have evolved a second line of defence to counter Stp by complexing PrrC with EcoprrI. Thus, when Stp physically inhibits EcoprrI, this complex is disrupted leading to the release of PrrC, which is an anticodon nuclease that targets tRNA-Lys. Depletion of this tRNA then leads to abortive infection. T4 can prevent abortive infection from occurring as Pnk and Rnl1 are able to repair the damage caused by PrrC. Thus, only T4 Pnk/Rnl1 mutants are aborted by PrrC and hence the name Prr (an abbreviation of pnk, rnl restriction) (Amitsur et al., 1987; Jabbar and Snyder, 1984). This is the opposite from what is observed from ΦCHAI8 and ToxIN$_{Pa}$ where deletion of the RNA repair protein homologs instead prevents abortive infection. Additionally, ΦCHAI8 does not have an Stp homolog and T4 itself is not aborted by ToxIN$_{Pa}$ suggesting a different mechanism for abortive infection.

The putative RNA repair system differs from that of T4 and other related RNA repair systems. In ΦCHAI8 and other *V5viruses*, the Pnk homolog is incomplete with only the kinase domain present. Instead there is a phosphatase located adjacent to it that accounts for the lack of the phosphatase domain. The phosphatase also doesn't match the phosphatase domain of Pnk but instead has homology to another nucleotide repair pathway protein involving double stranded DNA breaks. Interestingly, the yeast RNA repair system also shows differences to T4 in that all three activities i.e. kinase,

phosphatase and RNA ligase are all encoded on a single protein. Furthermore, the different domains can also be artificially expressed on three different polypeptides and still retain their respective activities. Homologs from T4 and mice has also been shown to be able to complement the different domains (Schwer et al., 2004, 2008). Thus it seems plausible that this RNA repair system can be found on three different proteins in ΦCHAI8. Whether these three proteins in ΦCHAI8 function as a RNA repair system is still to be determined biochemically, indeed a Pnk homolog of *Xanthomonas*, AvrRxo1-ORF1, which is both a Type III secretion effector and a Type II toxin is an important virulence factor (Han et al., 2015; Triplett et al., 2016). There is also an additional function of RnI1 which is in attachment of tail fibres to the baseplate (Rand and Gait, 1984; Wood and Henninger, 1969). Thus it is still unclear whether ToxIN$_{Pa}$ activation is caused by the putative RNA repair function.

The only other coliphage characterised with regards to Type III toxin-antitoxin mediated abortive infection is phage 2 active against the AbiQ system. Phage 2 escapes of the AbiQ system have mutations in a putative DNA polymerase. ΦCHAI8 also contains a homolog of this DNA polymerase, which raises some questions. It would be interesting to see if ΦCHAI8 is also aborted by AbiQ and, if aborted, whether it escapes through mutations in the DNA polymerase. Based on the amino acid sequence of the ΦCHAI8 and phage 2 DNA polymerase, it can be classified in DNA polymerase family A. This family of DNA polymerases are also known to have DNA repair capabilities due to exonuclease activity, and are involved in DNA repair caused by UV damage (Sharma and Smith, 1987). This is interesting for two reasons. Firstly, a theme is beginning to emerge whereby putative nucleotide repair pathways are implicated in Type III toxin-antitoxin mediated abortive infection. Secondly while there is no function yet assigned to the M1-23 protein of ΦM1 (**Chapter 4**) it has been shown to bind to UvrA, a protein involved in DNA damage repair caused by UV in bacteria. The ORF encoding M1-23 is also located between genes encoding an exonuclease and an endonuclease, perhaps implying an involvement with nucleotide metabolism/repair.

While ΦCHAI8 and phage 2 are both similar *V5viruses*, it is interesting to note that escapes of the two different Type III TA systems, ToxIN$_{Pa}$ and AbiQ have different mutations. This highlights again that activation of these Type III TA systems apparently use different mechanisms. T4 itself is strongly aborted by AbiQ but is completely

insensitive to ToxIN$_{Pa}$. However, it seems more than coincidental that the only coliphages capable of escaping ToxIN$_{Pa}$ and AbiQ are very similar so there could be a common factor to both phages that triggers activation of both ToxIN$_{Pa}$ and AbiQ. Further work will need to be done to further characterise the putative RNA repair system and the role it plays in ToxIN$_{Pa}$ escape.

# Chapter 6 Characterisation of phage escapes in different bacterial host backgrounds

## 6.1 Introduction

The Type III TA systems, ToxIN$_{Pa}$, TenpIN$_{Pl}$ and AbiQ, come from diverse bacterial backgrounds; ToxIN$_{Pa}$ from a plasmid of the potato pathogen *P. atrosepticum*, TenpIN$_{Pl}$ from the chromosome of the insect pathogen, *Ph. luminescens* and AbiQ from a plasmid originating from the dairy-associated Gram positive *Lactococcus lactis* (Blower et al., 2012b; Fineran et al., 2009; Samson et al., 2013). Even when transferred into a heterologous host (*E. coli*) they are still able to function as an abortive infection system against certain phages. Phages that are aborted also come from a wide variety of bacterial backgrounds and phages that are able to escape have widely different escape loci (Blower et al., 2017; Chen et al., 2017; Samson et al., 2013). This shows that activation of Type III TA cannot be specific to a single common protein that is phage derived, but rather a common process involving several different proteins from both the phage and the host harbouring the Type III TA system. One avenue that has not been investigated in detail is host specificity in response to Type III TA mediated abortive infection. A reason for this is that phages tend to be host specific as they have evolved to adsorb to specific host receptors and therefore it is difficult to isolate a phage that could infect widely different hosts. Furthermore, a narrow host range often means that the different hosts that a phage may infect would be highly similar and one might expect Type III TA system activation to be the same between similar hosts (for example between different common *E. coli* K12 strains such as DH5α and MG1655). This chapter examines the issue of host specificity in greater detail and discusses Type III TA activation of the same phage on hosts that are both similar or widely different.

## 6.2 Coliphages can escape Type III TA mediated Abi on one strain of *E. coli* but are unable to escape in another

The *E. coli* strain EPI300 has been previously used to test the effects of copy number on ToxIN$_{Pa}$ activation (Blower, 2009). The *E. coli* EPI300 strain was used as it contains a mutant *trfA* gene which codes for the TrfA protein required for the replication of

plasmids containing the *oriV* origin of replication. Thus EPI300 strains carrying a plasmid with this origin of replication (*e.g.* pCC1BAC) only carries one to two copies of the plasmid per cell. However, the copy number of pCC1BAC can be increased to 10-20 copies in the presence of TrfA that can be controlled by an inducible promoter. EPI300 was transformed with pTA46 (pBR322 containing $ToxIN_{Pa}$) which has a copy number of 15-20. This plasmid, pTA46 is pBR322 based and uses a different origin of replication and thus should be unaffected by the *trfA* mutation. When performing these experiments, it was noted that coliphage TB16 was aborted by DH5α carrying pTA46 and unable to escape. However, when TB16 infected EPI300 carrying pTA46, the phage was aborted but plaques of escape mutants could be seen at a frequency of $10^{-5}$ compared with the frame shift control plasmid, pTA47 (Blower, 2009). This led to the hypothesis that, although DH5α and EPI300 were similar K12 strains, there are small differences between the strains that could lead to a dissimilarity in $ToxIN_{Pa}$ activation. To examine whether these responses were limited to TB16, a selection of coliphages previously unable to evolve escapes on DH5α were plated on EPI300 carrying the same Type III TA containing plasmids (**Table 6.1**). These coliphages were additionally screened on DH5α to confirm that no escapes were observed. The data show that certain phages produced plaques on the EPI300 strain with a TA system which did not plaque on the same TA system on DH5α. This confirms that EPI300 must cause a differential TA activation compared to in DH5α, and for both the $ToxIN_{Pa}$ and $TenpIN_{Pl}$ systems. However, it should also be noted that not every phage that is aborted can evolve to escape in EPI300 (**Table 6.1**).

**Table 6.1** Certain coliphages produced plaques of escape mutants on EPI300 strains containing Type III TA systems that are not detectable in DH5α

| Phage | EOPs vs Type III TA plasmids | | | |
|---|---|---|---|---|
| | pTA46 ($ToxIN_{Pa}$) | | pFR2 ($TenpIN_{Pl}$) | |
| | DH5α | EPI300 | DH5α | EPI300 |
| ΦCHAI9 | ~1 | ~1 | $<1x10^{-8}$ | $1.7x10^{-4}$ |
| TRB18 | $<1x10^{-8}$ | $<1x10^{-8}$ | $<1x10^{-8}$ | $<1x10^{-8}$ |
| TRB19 | $<1x10^{-8}$ | $3.9x10^{-5}$ | ~1 | ~1 |
| F1 | $<1x10^{-8}$ | $<1x10^{-8}$ | $<1x10^{-8}$ | $<1x10^{-8}$ |
| F2 | $<1x10^{-8}$ | $6.2x10^{-5}$ | ~1 | ~1 |
| F4 | $<1x10^{-8}$ | $4.7x10^{-5}$ | $<1x10^{-8}$ | $4x10^{-9}$ |
| F5 | $<1x10^{-8}$ | $<1x10^{-8}$ | $<1x10^{-8}$ | $4.1x10^{-5}$ |
| F6 | $<1x10^{-8}$ | $<1x10^{-8}$ | $<1x10^{-8}$ | $<1x10^{-8}$ |
| F7 | $<1x10^{-8}$ | $8.6x10^{-5}$ | ~1 | ~1 |
| T4 | ~1 | ~1 | $<1x10^{-8}$ | $<1x10^{-8}$ |
| T6 | ~1 | ~1 | $<1x10^{-8}$ | $2.0x10^{-7}$ |

All phage lysates tested were “bulked” on DH5α with titres of at least $10^{10}$ pfu.ml$^{-1}$. EOPs were calculated by dividing the titre on the functioning Type III TA system with that on its frameshift control. Phages with EOPs of $<1x10^{-8}$ showed no detectable plaques on the functioning TA system despite testing a titre of at least $10^{10}$ pfu.ml$^{-1}$ on the frameshift control. Phages with EOPs of ~1 showed a similar number of plaques when compared to the frameshift control. N=1

Furthermore, there are cases where a phage (*e.g.* F5) that is aborted by one TA system and able to escape it on EPI300 cannot escape a different TA system in EPI300, highlighting again that there may be different activation mechanisms for the different TA systems. To confirm that phages had evolved escapes in EPI300, spontaneous escapes from three different phages were isolated. These escapes were shown to breed true as they had an EOP close to 1 compared to their frame shift control (**Table 6.2**). For phage F4 two escapes were picked representing escape from both the ToxIN$_{Pa}$ and TenpIN$_{Pl}$ system. F4E1 was selected on ToxIN$_{Pa}$ as it was insensitive to ToxIN$_{Pa}$. However, it still remained sensitive to the TenpIN$_{Pl}$, system although not aborted as strongly. F4E2 on the other hand was selected on the TenpIN$_{Pl}$ system as it was insensitive to that system but still highly sensitive to ToxIN$_{Pa}$. This suggests that, although the activation of the Type III TA systems appeared to be unlinked, there may be factors that affect both of them, albeit to different degrees.

**Table 6.2** Escapes of coliphages isolated on EPI300 carrying Type III TA systems

| Phage | TA system | EOP |
|---|---|---|
| ΦCHAI9 | pFR2 | $1.7x10^{-4}$ |
| ΦCHAI9-E1 | pFR2 | 0.88 |
| F4 | pTA46 | $4.7x10^{-5}$ |
| | pFR2 | $4.0x10^{-9}$ |
| F4-E1 | pTA46 | 0.53 |
| | pFR2 | $4.7x10^{-3}$ |
| F4-E2 | pTA46 | $2.2x10^{-5}$ |
| | pFR2 | 1.2 |
| T6 | pFR2 | $2.0x10^{-7}$ |
| T6-E1 | pFR2 | 0.06 |

ΦCHAI9-E1, escape of ΦCHAI9 selected on pFR2

F4-E1, escape of F4 selected on pTA46

F4-E2, escape of F4 selected on pFR2

T6-E1, escape of T6 selected on pFR2

## 6.3 Phage T6 can escape the TenpIN$_{PI}$ system via mutations in a transcriptional inhibitor in EPI300

As it was not possible to sequence all the phages and their escapes, one phage was picked for further characterisation. Phage T6 was chosen as, being one of the T-even phages, it was similar to the best characterised phage T4 and thus it was likely that any mutations found in escapes may be in a protein that had been studied before. There also had not been any coliphage escape of the TenpIN$_{PI}$ system yet characterised. Thus two escapes selected on TenpIN$_{PI}$ as well as the wild type T6 genomic DNA was sequenced. Data from whole genome sequencing revealed that both escapes had mutations in the *alc* gene which codes for Alc, a transcriptional inhibitor (**Table 6.3**). Six more escapes were subsequently isolated and their *alc* genes sequenced. Two of these new escape mutants also had mutations in *alc* but the remaining four having no mutations in this gene implying that there is an additional escape locus in these T6 escapes.

The function of the Alc protein was discovered originally through work using mutants of the T4 phage that incorporate cytosine into their DNA, during replication. The wild type T4 phage (and the other T-evens, T2 and T6) have glucosylated hydroxymethylcytosine in place of cytosine (Lehman and Pratt, 1960). These T4 mutants that instead utilise cytosine are unable to produce progeny and this was thought to be because they could not produce any of their late RNA or proteins (Kutter et al., 1975). However, mutants were isolated that had mutations in a single gene that allowed late RNA and proteins to be produced, even with cytosine incorporated into the genome (Snyder et al., 1976). This gene was termed “allows transcription of late protein genes from T4 cytosine containing DNA” or *alc* for short. However, this name misrepresents its function as subsequent research showed that it blocks cytosine containing DNA of the host as well as other phages (Kutter et al., 1981; Pearson and Snyder, 1980). The predicted role of Alc is that it is used by T4 (and other T-even phages) to identify and differentiate its genomic DNA from host DNA to specifically halt host transcription while allowing transcription of viral genes with modified cytosine-containing genomes. Alc is thought to interact with an actively transcribing RNA polymerase and is important in the switch from host to phage metabolism (Drivdahl and Kutter, 1990; Severinov et al., 1994). Furthermore, another gene designated *unf*

which encoded a protein involved in unfolding the bacterial nucleoid, was shown to be the same gene as *alc* (Snyder and Jorissen, 1986).

Both the Alc proteins in T4 and T6 are 167 amino acids long with only four different residues between them (**Figure 6.1**). Alignments of mutations that have been found previously in T4 mutant with an $Alc^-$ phenotype (Snyder and Jorissen, 1986) and T6 Alc escapes of $TenpIN_{PI}$ are shown in **Figure 6.1**. None of the T4 or T6 mutations are the same, with the T4 mutations located more N-proximal than the T6 mutations. T6 *alc* mutations are very C-terminal, the T6E2 mutant only replaces eight amino acids with three different amino acids. This is vaguely reminiscent of the M1-23 protein escape of which versions have only minor alterations to the C-terminal domain. However, completely abolishing Alc activity may not be necessary to escape $TenpIN_{PI}$ and perhaps only a slight disruption of protein activity is sufficient.

```
T4 wt       MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITGLKKDLAKQVKRKTTIKEYRNRRLE   60
Alc10       MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITGLKKDLAKQVKRKTTIKEYRNRRLE   60
Alc536      MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITGLKKDLAKQVKRKTTIKEYRNRRLE   60
Alc2        MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITGLKKDLAKQVKRKTTIKEYRNRRLE   60
AlcLJ7      MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITGLKKDLAKQVKRKTTIKEYRNRRLE   60
T6 wt       MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITDLKKDLAKQVKRKTTIKEYRNRRLE   60
T6E1        MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITDLKKDLAKQVKRKTTIKEYRNRRLE   60
T6E2        MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITDLKKDLAKQVKRKTTIKEYRNRRLE   60
T6E3        MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITDLKKDLAKQVKRKTTIKEYRNRRLE   60
T6E5        MDLQLITTEMVVEAYGDTTDGISVFKGNRRVGYITDLKKDLAKQVKRKTTIKEYRNRRLE   60

T4 wt       QARDMLPDAVEEMKVFLENQLAKYDCEVFINQTQPNVHINSCKCYIIVNPLTGKHRLGIS   120
Alc10       QARDMLPDAVEEMKVFLENQLAKYDCEVFINQTQPNVHINSCKYYIIVNPLTGKHRLGIS   120
Alc536      QARDMLPDAVEEMKVFLENQLAKYDCEVFINQTQPNVHINSCKCYIIVNPLTGKRRLGIS   120
Alc2        QARDMLPDAVEEMKVFLENQLAKYDCEVFINQTQPNVHINSCKCYIIVNPLTGKHRLDIS   120
AlcLJ7      QARDMLPDAVEEMKVFLENQLAKYDCEVFINQTQPNVHINSCKCYIIVNPLTGKHRLGIS   120
T6 wt       QARDMLPDAVEEMKVFLENQLAKYDCDVFINQTQPNVHINSCKCYIIVNPLTGKHRLGIS   120
T6E1        QARDMLPDAVEEMKVFLENQLAKYDCDVFINQTQPNVHINSCKCYIIVNPLTGKHRLGIS   120
T6E2        QARDMLPDAVEEMKVFLENQLAKYDCDVFINQTQPNVHINSCKCYIIVNPLTGKHRLGIS   120
T6E3        QARDMLPDAVEEMKVFLENQLAKYDCDVFINQTQPNVHINSCKCYIIVNPLTGKHRLGIS   120
T6E5        QARDMLPDAVEEMKVFLENQLAKYDCDVFINQTQPNVHINSCKCYIIVNPLTGKHRLGIS   120

T4 wt       NPNRSASDMAEDVEACFKISKSPAEHHILINGLSQDDIVEVIKTLCM--    167
Alc10       NPNRSASDMAEDVEACFKISKSPAEHHILINGLSQDDIVEVIKTLCM--    167
Alc536      NPNRSASDMAEDVEACFKISKSPAEHHILINGLSQDDIVEVIKTLCM--    167
Alc2        NPNRSASDMAEDVEACFKISKSPAEHHILINGLSQDDIVEVIKTLCM--    167
AlcLJ7      ILVSKS----------------------------------------------    126
T6 wt       NPNRSASDMAEDVEACFKISKSPAEHHILINGLSQDDIIEVIKTLCN--    167
T6E1        NPNRSASDMAEDAEACFKISKSPAEHHILINGLSQDDIIEVIKTLCN--    167
T6E2        NPNRSASDMAEDVEACFKISKSPAEHHILINGLSQDDIIGGY-------    162
T6E3        NPNRSASDTAEDVEACFKISKSPAEHHILINGLSQDDIIEVIKTLCN--    167
T6E5        NPNRSASDMAEDVEACFKISKSPAEHHILINGLSQDDIIRRLLKLYAIK    169
```

**Figure 6.1** Alc alignments of T4, T6 and mutants

In green are residues that are different between T4 and T6 Alc proteins. Residues highlighted in purple and orange are differences in residues caused by mutations in the *alc* gene in T4 and T6 respectively.

**Table 6.3** ΦT6 escape mutants

| Isolation host | Phage | EOP on, EPI300 (pFR2) [TenpIN$_{PI}$] | Mutation in *alc* | Effect on protein |
|---|---|---|---|---|
| EPI300* | T6 | 2.0 x $10^{-7}$ | None | None |
| EPI300, (pFR2) [TenpIN$_{PI}$] | T6-E1 | 0.056 | T398C | V133A |
| | T6-E2 | 0.042 | -478G | Frameshift |
| | T6-E3 | 0.21 | T386C | M129T |
| | T6-E4 | 0.38 | No mutation | None |
| | T6-E5 | 0.050 | -477AG | Frameshift |
| | T6-E6 | 0.047 | No mutation | None |
| | T6-E7 | 0.039 | No mutation | None |
| | T6-E8 | 0.19 | No mutation | None |

## 6.4 Differences between DH5α and EPI300

There is now clear evidence that host related factors can play a role in Type III TA activation. Nonetheless it is still curious that both DH5α and EPI300 carry the same plasmids and yet exhibit different phenotypes in relation to abortive infection. The disparity then must be the result of different mutations between the two K12 strains. Both DH5α and EPI300 are highly modified *E. coli* strains designed for specific roles. DH5α is one of the most used strains for gene cloning in molecular biology while EPI300 was designed to control the copy number of OriV-based plasmids. A closer analysis of the known genotype of DH5α and EPI300 shows that they share some of the same mutations which presumably play no role in the different Abi responses seen here (**Table 6.4**). However, there are multiple mutations not shared between the two strains (**Table 6.5 and 6.6**). Inspection of each of these mutations does not point to any obvious reasons for the difference in phenotype with regards to abortive infection. The only mutation that perhaps may have a role is *trfA* in EPI300, which prevents the replication of RK2 based plasmids. This mutation causes plasmids with the OriV origin of replication to be maintained at a single copy. However, there are no reports of *trfA* affecting the copy number of pBR322, so the cause of a differential response in the Abi system is not obvious. Nonetheless, copy number of plasmids has been shown to be important for ToxIN$_{Pa}$ activation (Blower, 2009) so further work will be required to assess the copy number of pBR322 in EPI300 and DH5α. It is also possible that there are other mutational differences between the strains that are unknown and play a role in Type III TA activation.

**Table 6.4**. Mutations shared between DH5α and EPI300

| Abbreviation | Consequence |
|---|---|
| $F^-$ | Does not carry F plasmid |
| *endA1* | Non-functional endonuclease I which prevents non-specific digestion |
| *recA1* | Non-functional RecA DNA repair protein. Prevents unwanted DNA recombination in cloning |
| *nupG* | Nucleotide transporter deletion allows uptake of large plasmids |
| Φ80*dlacZΔM15* | Carries defective Φ80 prophage with partial deletion of *lacZ* gene (used for blue and white screening on XGal) |
| $\lambda^-$ | Non lysogen (λ) |

**Table 6.5** Mutations in DH5α not present in EPI300

| Abbreviation | Consequence |
|---|---|
| *glnV44* | Suppression of amber (UAG) stop codons which is required for certain phages to replicate |
| *thi-1* | Thiamine auxotroph |
| *relA* | ppGpp synthesis during the stringent response. Required for RNA synthesis in the absence of protein synthesis |
| *gyrA*96 | Mutation in DNA gyrase leading to nalidixic acid resistance |
| *deoR* | Mutation in DeoR transcriptional repressor. Allows uptake of large plasmids (same result as *nupG*) |
| *purB*20 | Mutation in adenylosuccinate lyase leading to slow growth in M9 |
| (*lacZYA-argF*)U169 | Deletion of lac operon and *argF* |
| *hsdR*17($r_K^- m_K^+$) | Leads to loss of function of the restriction endonuclease from RM system *EcoKI* |

**Table 6.6** Mutations in EPI300 not present in DH5α

| Abbreviation | Consequence |
|---|---|
| *mcrA*Δ(*mrr-hsdRMS-mcrBC*) | Complete deletion of RM system EcoKI |
| Δ*lac*X74 | Deletion of entire lac operon |
| *ara*D139 | Prevents arabinose metabolism by mutation in L-ribulose-phosphate 4-epimerase |
| Δ(*ara*,*leu*)7697 | Deletion of an operon involved in the biosynthesis of leucine from valine. Auxotrophic for leucine. |
| *galU, galK* | Mutation prevents metabolism of galactose |
| *rpsL* | Mutation in ribosomal protein S12 that leads to streptomycin resistance |
| *trfA* | Gene encodes TrfA protein that is required for the replication of RK2 based plasmids (OriV origin of replications) |
| *dhfr* | Encodes dihydrofolate reductase which confers trimethoprim resistance |

## 6.5 ΦOT8s

### 6.5.1 ΦOT8s is a LamB dependent *Tevenvirinae Serratia* phage

The *E. coli* DH5α and EPI300 experiments have shown that host related factors play a role in Type III TA activation. However, another phage had already been isolated that is capable of infecting two even more distantly related hosts that was sensitive to Type III TA systems (Blower, 2009). This phage, ΦOT8s, provided a further opportunity to investigate host related factors involved in Type III TA activation. ΦOT8s was isolated as a phage forming plaques on *Serratia* strain 39006 carrying a plasmid expressing the $LamB_{E.coli}$ protein (pMUT13) (Blower, 2009). It is unable to infect *Serratia* without the $LamB_{E.coli}$ protein but can infect strains of *E. coli* that naturally express the LamB protein (Blower, 2009). Sequencing results prior to this study revealed that ΦOT8s belongs to the *Tevenvirinae* subfamily (often referred to as T4-like phages) and it shared high sequence homology with other *E. coli* phages of the same subfamily (**Figure 6.2**). ΦOT8s is closely related to Rb69 (accession number AY303349) sharing 89.8% nucleotide identity and thus belongs to the same genus *i.e. Rb69virus*. ΦOT8s shares only 61.7% nucleotide identity with phage T4. The sequence of ΦOT8s was annotated again during this study as previous annotation was several years ago. The new annotation of ΦOT8s predicts 271 ORFs with 2 tRNAs. A genomic map of ΦOT8s is also shown in **Figure 6.3**.

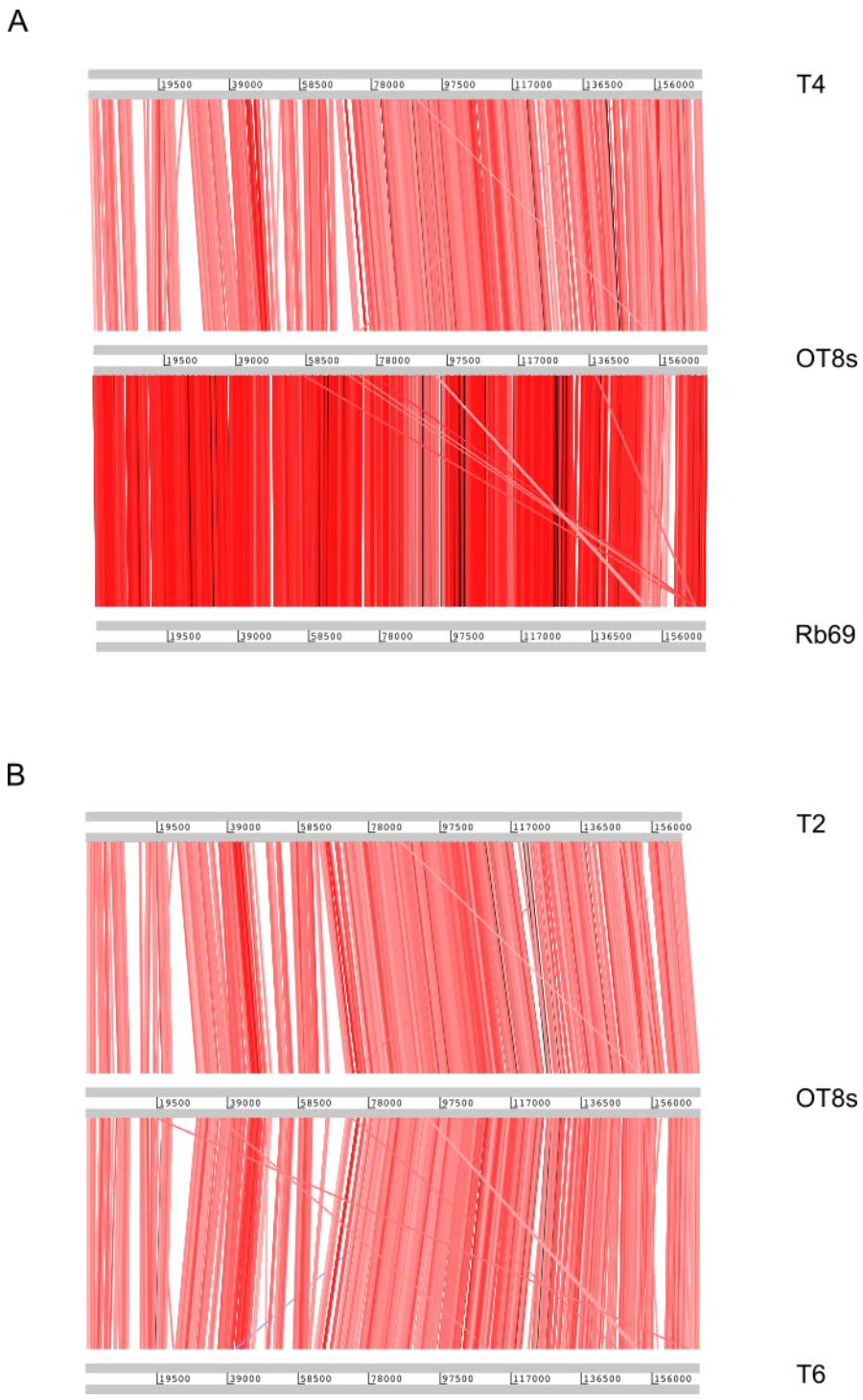



**Figure 6.2** Genome wide comparisons of ΦOT8s and related *Tevenvirinae*

**A**. ΦOT8s shows higher homology to Rb69 than T4. **B**. ΦOT8s compared to T2 and T6. Comparison files were generated using a tblastx comparison with a $1x10^{-4}$ cut off value.

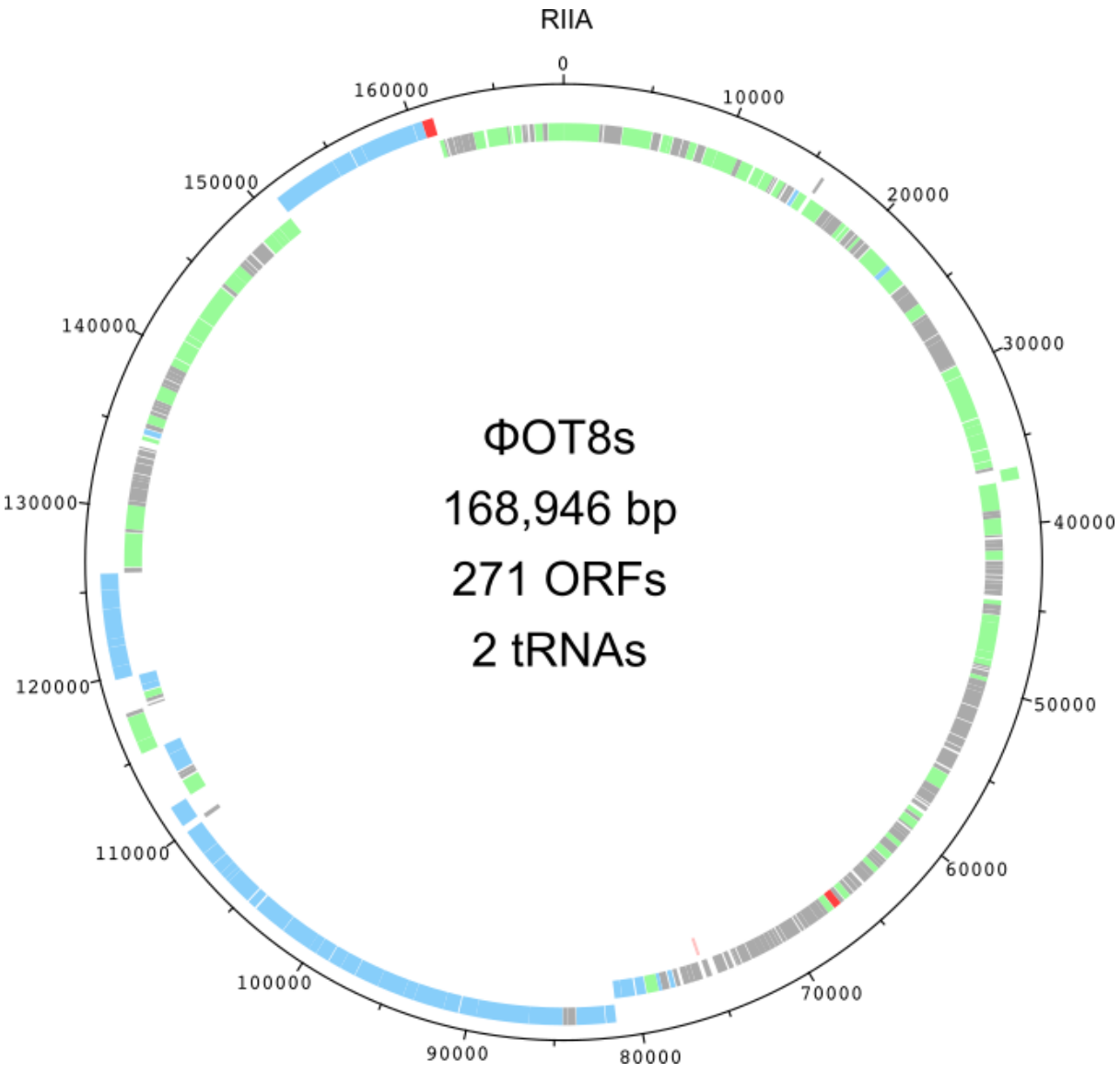



**Figure 6.3** Genomic map of ΦOT8s

The genome has been orientated to begin at RIIA as with other *Tevenvirinae*. Boxes in green, blue and red signify ORFs encoding proteins associated with metabolism, virion structure and lysis. Grey boxes signify ORFs that encode hypothetical proteins. The pink box in the inner most ring shows the location of the tRNA (~75,000 bp).

The *Tevenvirinae* subfamily is named after the T-even phages and includes a wide range of phages that share several conserved proteins referred to as the core genome (Comeau et al., 2007). The receptors for *Tevenvirinae* phages tend to be outer membrane proteins with a few using LPS as a secondary receptor (Trojet et al., 2011). Several of the *Tevenvirinae* phages, including T2 and T6, use an adhesion protein called Gp38 to recognise their outer membrane protein receptors (Trojet et al., 2011). In these phages assembly of the tail is completed by cleavage of the C terminal end of Gp37 the distal tail fibre. The adhesion protein Gp38 then replaces the cleaved region of Gp37. Not every *Tevenviriniae* member constructs its tail fibres in this way and many of the host recognition adhesion proteins for these phages have not yet been identified. T4 is one of the phages that deviates from this tail fibre configuration. The C terminal region of the Gp37 of T4 is not cleaved and resembles the Stf protein of phage λ, a very different phage (George et al., 1983). Instead Gp37 by itself is able to recognise the T4 receptors, OmpC and LPS (Montag et al., 1990). Mutations of Gp37 can also lead to an expanded host range in T4 (Tétart et al., 1996). Furthermore, the Gp38 protein in T4 does not share any homology with the Gp38 of phages that use it as an adhesion protein. Instead Gp38 and another protein in T4 (Gp57) act as chaperones that aid the folding of the Gp37 protein (Hashemolhosseini et al., 1996). Additionally the homolog of Gp38, (Tfa in λ) as well as the C-terminal region of Stf is able to complement defects in the Gp37 and Gp38 protein of T4 (Montag et al., 1989). Analysis of the Gp37 and Gp38 homologs of ΦOT8s revealed that they shared homology with those of T4 rather than Rb69 (**Figure 6.4 A**). Comparisons with the λ proteins, Stf and Tfa also show that the C-terminal region of Stf and Tfa have amino acid sequence homology with both the C-terminal of Gp37 and Gp38 of ΦOT8s. λ is well known to use LamB as a receptor which is recognised by the J protein, forming part of the tail spike (Wang et al., 1998). However, this protein has no homology with any protein of ΦOT8s (**Figure 6.4 B**). With regards to the Stf protein, the common wild type lab strain of λ has a frame shift mutation in this protein rendering it non-functional (Hendrix and Duda, 1992). The original λ isolate on the other hand has an intact Stf protein that binds to OmpC and thus has a broader host range (Hendrix and Duda, 1992). The mutant λ that is commonly used today is thought to have been picked (by accident) from the original lysate as plaques produced by the original strain were smaller which was thought to be as a result of the tail fibres causing a slower diffusion in the agar (Casjens and Hendrix, 2015). The frame shift in the Stf protein has caused

it to be split into two proteins originally annotated as *orf401* and *orf314* (Sanger et al., 1982). A schematic showing the different tail fibre/tail spike configurations of *Tevenvirinae* and λ is shown in **Figure 6.5**. The receptor binding region that recognises OmpC or LPS within the C-terminal tail of T4 Gp37 has been shown to be at residues 907 to 996 (Bartual et al., 2010; Montag et al., 1990). Comparisons with highly similar T4-like phage Gp37 proteins from phages TuIa, TuIb or the Stf protein from λ which recognise either OmpF or OmpC (Gp37 of T4 also recognises LPS) show that the receptor binding region is located in a variable area of the proteins (Montag et al., 1990). Alignments with the Gp37 protein of ΦOT8s also shows variability where the receptor binding region is predicted to be (**Figure 6.6**). The Gp37 protein of ΦOT8s also shows higher similarity with the Gp37 proteins of λ, TuIa and TuIb than T4, containing a predicted eighth metal binding site compared to seven in T4 (Bartual et al., 2010).

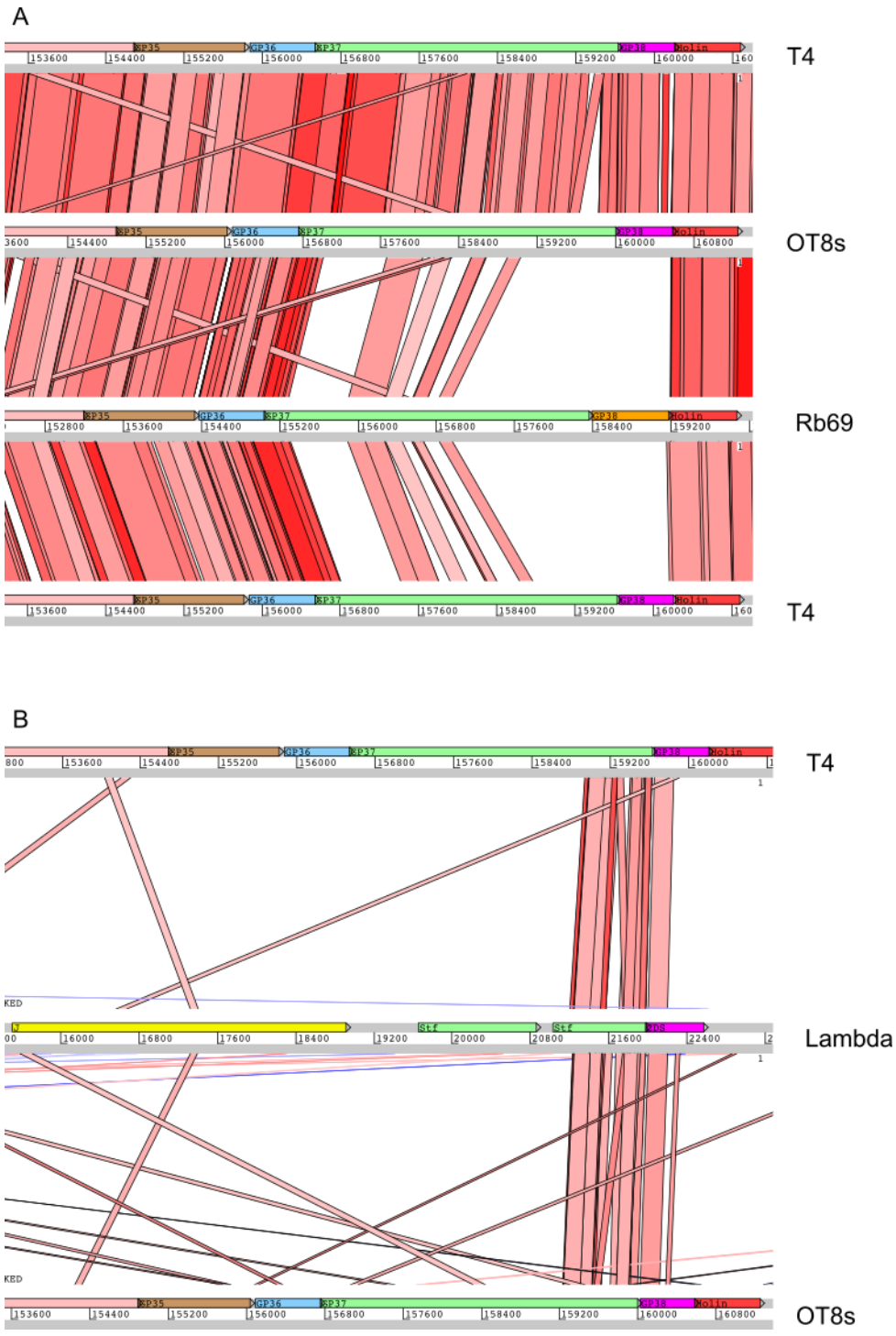



**Figure 6.4** Alignments of proteins involved in tail fibre construction

**A**. ΦOT8s shows high amino acid sequence homology to the tail fibre proteins of T4. Rb69 additionally shows no homology in its Gp38 protein (orange) to the Gp38 protein (purple) of ΦOT8s and T4. There is also poor homology between the Gp37 protein

(green) of Rb69 against ΦOT8s and T4. **B**. The C-terminal region of Gp37 of ΦOT8s and T4 show homology with the C-terminal end of Stf (green) and Gp38 (purple). The λ sequence used is the standard λ wild type strain which has a frame shift in the middle of the Stf protein causing it to be split into two proteins (J02459). Moreover, the J protein of λ (yellow) shows no homology to any ΦOT8s or T4 proteins.

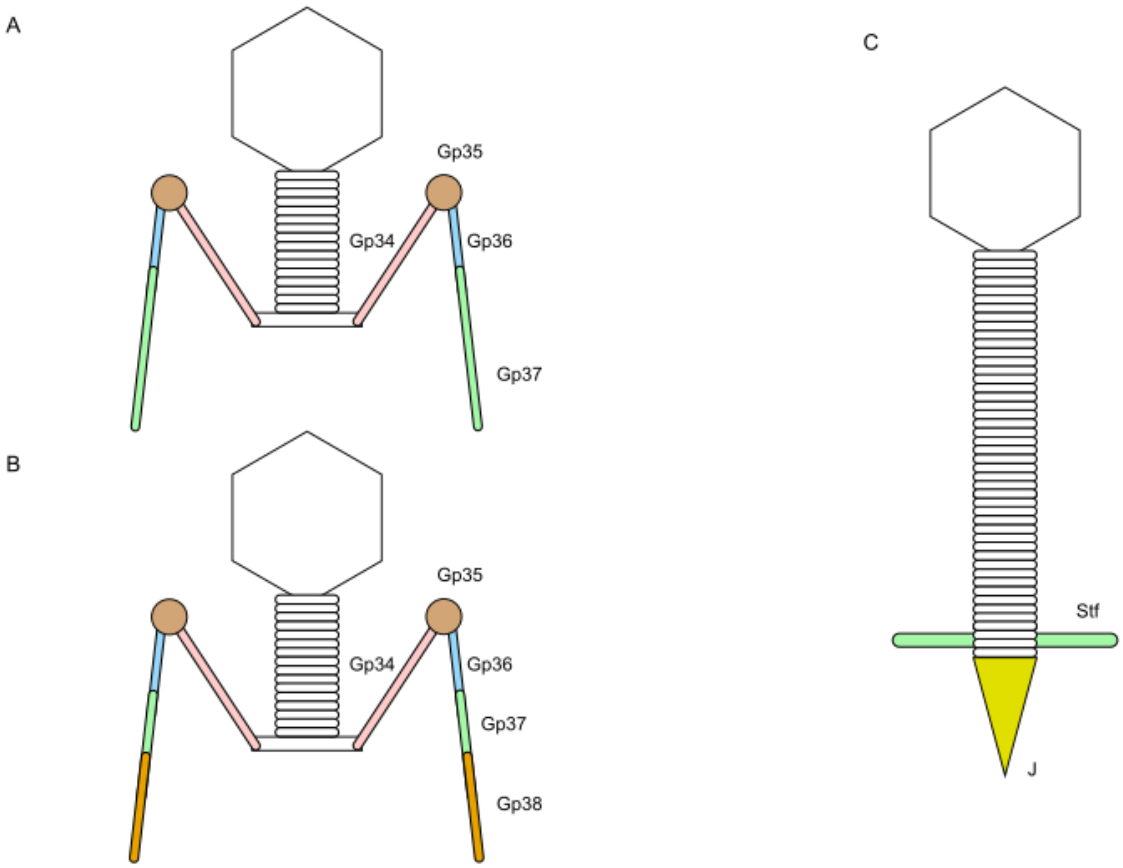



**Figure 6.5** Tail fibre organisation of *Tevenvirininae* and λ

Tail fibre organisation for **A**. ΦOT8s and T4, **B**. Majority of *Tevenvirinae* including T2 and T6 and **C**. λ. In **A** and **B** tail fibre construction is similar apart from the tail end where Gp37 is not cleaved in **A**. In **B** Gp37 is proteolytically cleaved with the cleaved region replaced with another protein Gp38. **C**. The original λ isolate had side tail fibres made from the Stf protein which has homology with the C-terminal ends of the Gp37 protein of ΦOT8s and T4. The LamB recognition protein in λ is protein J which forms part of the tail spike.

```
T4          811  SSYPIGAPIPWPSDSVPAGFALMEGQTFDKSAYPKLAVAYPSGVIPDMRGQTIKGKP-SG 869
TuIa              SSYPIGAPIPWPTDTPPNGYALMEGQTFDTRAYPKLAAAYPSGTIPDMRGQTIKGKP-SG
TuIb              SSYPIGAPIPWPTDTPPDGYAIMEGQTFDKSAYPKLAMAYPSGVIPDMRGQTIKGKP-SG
Lambda      528  SAFPAGAPIPWPSDIVPSGYVLMQGQAFDKSAYPKLAVAYPSGVLPDMRGWTIKGKPASG 587
OT8s        824  SSYPIGAPIPWPTDTPPEGYAIMEGQTFDADLYPKLAAVYPSGTLPDMRGQTIKGKP-SG 882
                 *::* *******:*  * *:.:*:**:**   ***** .****.:***** ****** **

T4          870  RAVLSAEADGVKAHSHSASASSTDLGTKTTSSFDYGT----------KGTNSTGGHTHSG 919
TuIa              RAVLSTEADGVKSHTHGASASNTDLGTKTTSSFDYGTKTTSSFDYGTKTSNTTGNHNHTV
TuIb              RAVLSAEADGVKAHSHSASASSTDLGTKTTSSFDYGTKTTSSFDYGTKTSNTTGNHSHTV
Lambda      588  RAVLSQEQDGIKSHTHSASASGTDLGTKTTSSFDYGTKTTGSFDYGTKSTNNTGAHAHSL 647
OT8s        883  RAVLSTEADGVKSHSHSASASSTDLGTKTTSSFDYGSKTTSSFDYGTKTSNTTGNHNHTV 942
                 ***** * **:*:*:*.****.**************:          * :*.** * *:

T4          920  SGSTSTNGEHSHYI---------------EAWNGTGVGGNKMSSYAISYR--A---GGSN 959
TuIa              SGTTSSAGAHQHARSGPQLSNGISTNIFPDGYSDVGTNYNSKFSGTVIGS--SVPCIIGK
TuIb              SGTTSSAGAHQHARSGPQIQAGIPNNIFYDGYNSVGANANAKITGSVSGA--TAGSNIAK
Lambda      648  SGSTGAAGAHAHTSGLRMN---------SSGWSQYGTATITGSLSTVKGTSTQGIAYLSK 698
OT8s        943  SGNTSSAGAHQHARSGPQVQGGIPTSIFYDGYNSAGPNGNAKFTATVSGA--TASSNMAK 1000
                 **.*.: * * *                  ..:.  *         ::           .:

T4          960  TNAAGNHSHTFSFGTSSA---------GDHSHSVGIGAHTHTVAIGSHGHTITVNSTGNT 1010
TuIa              TSNDGAHTHTWSGTTSTTGNHAHTVGIGAHTHTVGIGAHTHTVAIGSHGHTITVNATGNT
TuIb              TSSDGAHTHTWSGTTSTTGNHAHTVGIGAHTHTVGIGAHTHTVAIGSHGHTITVNSTGNT
Lambda      699  TDSQGSHSHSLSGTAVSAGAHAHTVGIGAHQHPVVIGAHAHSFSIGSHGHTITVNAAGNA 758
OT8s        1001 TSSDGAHTHSWSGTTSTTGNHAHTVGIGAHTHSVGIGAHSHTVAIGSHGHTITVNATGNT 1060
                 *.  * *:*: *  : ::         * * * * ****:*:.:***********::**:

T4          1011 ENTVKNIAFNYIVRLA    1026
TuIa             ENTVKNIAFNYIVRLA
TuIb             ENTVKNIAFNYIVRLA
Lambda      759  ENTVKNIAFNYIVRLA    774
OT8s        1061 ENTVKNVAFNYIVRLA    1076
                 ******:*********
```

**Figure 6.6** Alignments of the C-terminal regions of Gp37 from T4, TuIa, TuIb, λ and ΦOT8s

There is only a partial sequence of the Gp37 protein known for TuIa and TuIb thus the residue locations were not added. However, it is predicted to be of a similar size to the Gp37 protein of T4. The residues shaded in grey are predicted to be involved in the recognition of OmpC and LPS in T4. This is in a variable region with residues outside of this area being more conserved. In bold are the metal binding domains of which there are seven in T4 and eight in the other phages shown. Known receptors for the Gp37 proteins of the phages are T4 (OmpC and LPS), TuIa (OmpF), TuIb (OmpC only), λ (OmpC) and ΦOT8s (LamB).

### 6.4.2 ΦOT8s escapes have mutations in *motA* and *asiA*

ΦOT8s was previously found to be aborted by the ToxIN$_{Pa}$ system in both *Serratia* carrying pMUT13 and *E. coli* DH5α, and able to evolve escapes (Blower, 2009). Thus it provided a rare opportunity in which Type III TA activation from the same phage could be studied in different bacteria. Sixteen ToxIN$_{Pa}$ escapes were selected, eight from *Serratia* 39006 (pMUT13, pTA46) and the other eight from *E. coli* DH5α (pTA46) (**Table 6.7**). All escapes selected on *Serratia* (pMUT13, pTA46) were also shown to be insensitive to ToxIN$_{Pa}$ - carrying DH5α (pTA46). On the other hand, not all escapes isolated on DH5α were insensitive to *Serratia* (pMUT13) carrying ToxIN$_{Pa}$. ΦOT8s-E14 was still aborted by *Serratia* (pMUT13, pTA46) even though it had already been exposed to the ToxIN$_{Pa}$ system in DH5α (pTA46). This suggests that there may be multiple mutations involved where certain mutations can lead to insensitivity to two hosts while some only lead to insensitivity to one.

Prior to this study several ΦOT8s escapes were isolated (Blower, 2009). Several of these escapes were sent for whole genome sequencing and results showed that they all had mutations in a single gene that encodes MotA. Sequencing of the *motA* gene in the newly isolated ΦOT8s escapes ΦOT8s-E1-16 revealed that six of the escapes had *motA* mutations (**Table 6.8**). One of the newly isolated escapes that did not have a *motA* mutation (ΦOT8s-E1) was sent for whole genome sequencing, alongside a wild type ΦOT8s as a control. The results from the sequencing revealed mutations in the *asiA* gene. Subsequent sequencing of genomes of escape mutants without *motA* mutations also revealed that some of the escapes also had *asiA* mutations as well (**Table 6.8**). However, there were several escapes in which there were no mutations in *motA* and *asiA* conferring that for some escapes, there are other mutations not yet identified. For the *motA* mutants three different mutations were observed. Two of the mutations were substitutions in base G404, which changed to T or A leading to Arg102 being substituted by lysine or histidine respectively. The remaining *motA* mutation is a substitution of T110A leading to a valine to alanine conversion. All *asiA* mutations were identical and were caused by a single adenine insertion at position 127 leading to a frameshift mutation.

**Table 6.7** ToxIN$_{Pa}$ escapes of ΦOT8s

| Isolation host | Phage | EOPs<br>39006 (pMUT13, pTA46) | DH5α (pTA46) |
|---|---|---|---|
| *Serratia* ATCC 39006 (pMUT13) | ΦOT8s (wt) | $2.2 \times 10^{-5}$ | $2.0 \times 10^{-3}$ |
| *Serratia* ATCC 39006 (pMUT13, pTA46) | ΦOT8s-E1 | 0.18 | 0.78 |
| | ΦOT8s-E2 | 0.25 | 0.92 |
| | ΦOT8s-E3 | 0.26 | 0.84 |
| | ΦOT8s-E4 | 0.71 | 0.92 |
| | ΦOT8s-E5 | 0.70 | 1.2 |
| | ΦOT8s-E6 | 0.59 | 0.57 |
| | ΦOT8s-E7 | 0.47 | 1.1 |
| | ΦOT8s-E8 | 0.38 | 0.56 |
| *E. coli* DH5α (pTA46) | ΦOT8s-E9 | 0.43 | 0.73 |
| | ΦOT8s-E10 | 0.061 | 0.62 |
| | ΦOT8s-E11 | 0.05 | 0.38 |
| | ΦOT8s-E12 | 0.74 | 0.86 |
| | ΦOT8s-E13 | 0.30 | 0.05 |
| | ΦOT8s-E14 | $1.9 \times 10^{-5}$ | 0.12 |
| | ΦOT8s-E15 | 1.2 | 1.14 |
| | ΦOT8s-E16 | 0.03 | 0.41 |

pMUT13 – plasmid expressing LamB$_{E.coli}$

pTA46 – pBR322 expressing the ToxIN$_{Pa}$ system

EOPs are calculated by dividing the titre of phage on the strain carrying the function Type III TA system with the titre on the strain carrying the corresponding frameshift.

**Table 6.8** *motA* and *asiA* mutations in ΦOT8s escapes

| Phage | *motA* mutation | *asiA* mutation | Effect on protein |
|---|---|---|---|
| ΦOT8s (wt) | None | None | None |
| | | | |
| ΦOT8s-E1 | None | -127A | Frameshift |
| ΦOT8s-E2 | G404T | None | R135L |
| ΦOT8s-E3 | None | -127A | Frameshift |
| ΦOT8s-E4 | None | -127A | Frameshift |
| ΦOT8s-E5 | G404A | ND | R135H |
| ΦOT8s-E6 | G404T | None | R135L |
| ΦOT8s-E7 | G404T | None | R135L |
| ΦOT8s-E8 | None | -127A | Frameshift |
| | | | |
| ΦOT8s-E9 | T110A | ND | V37D |
| ΦOT8s-E10 | None | -127A | Frameshift |
| ΦOT8s-E11 | None | None | None |
| ΦOT8s-E12 | None | None | None |
| ΦOT8s-E13 | None | None | None |
| ΦOT8s-E14 | None | None | None |
| ΦOT8s-E15 | G404A | ND | R135H |
| ΦOT8s-E16 | None | None | None |

ND – not determined *i.e.* not sequenced

"-" followed by a number and a letter represents an insertion at that position

Genes for the *Tevenvirinae* are expressed at different stages of the phage lifecycle and are divided into three groups called early, middle and late, expressed sequentially (Miller et al., 2003). The promoters of early genes of *Tevenvirinae* closely resemble those of their hosts and have higher affinities thus redirects the host RNAP to preferentially cause the transcription of early phage genes. The switch from early to middle genes is controlled by two early proteins, MotA and AsiA through a process called σ appropriation (Ouhammouch et al., 1995). The promoter sequences of the middle genes differ from those of early genes as they contain a different -35 sequence which is a conserved 9 base pair sequence called the MotA box (Cicero et al., 1998). This sequence is recognised and bound by a single MotA protein (Cicero et al., 1998). The MotA protein is in two domains where the N-terminal domain binds sigma 70 while the C-terminal domain contains the DNA binding domain that recognises and binds to the MotA sequence (Li et al.; Pande et al., 2002). As MotA redirects the RNAP complex through its σ 70 subunit, AsiA prevents the transcription of both the host and early gene promoters. In the absence of phage infection, σ 70 binds to the -35 promoter regions though a conserved region called 4.2 (Campbell et al., 2002; Paget and Helmann, 2003). AsiA blocks the 4.2 interaction with the -35 promoter region by binding to 4.2 (Colland et al., 1998b). This then prevents the transcription of all early and host promoters that utilise this -35 sequence. However, as mentioned above, phage -35 middle gene promoters can still be recognised due to the interactions with MotA. **Figure 6.7** shows a schematic of how the different proteins interact during sigma appropriation in T4.

The amino acid sequence of T4 MotA and ΦOT8s and a representative of the different three mutations have been aligned (**Figure 6.8**). One of the MotA mutants has a mutation in the N-terminal domain although it is not located in any of the T4 MotA residues that are predicted to interact with the σ 70 unit (Bonocora et al.). However, previous experiments with the mutant MotA protein have also identified residues important in transcription that are not in the predicted interaction site (Finnin et al., 1997). The C-terminal mutations in the MotA protein are far more common and are all substitution mutations of arginine 135. Structural analysis of the T4 C-terminal domain of MotA predicts that arginine 135 is one of the residues involved in DNA binding (Li et al.).

The *asiA* mutations for ΦOT8s mutants escaping $ToxIN_{Pa}$ are all caused by an insertion of an adenine at position 127. These frameshift mutants generate a premature stop codon and effectively changes the last 46 amino acids of the wild type AsiA protein (**Figure 6.9**). The role of the C-terminal region in T4 AsiA is somewhat controversial. One study has shown that deletion of the last forty amino acids of T4 AsiA had no effect on transcription (Pal et al., 2003). Another showed that single substitution mutations in residues within the last forty amino acids deleted in the aforementioned study does indeed have an effect on transcription, and additionally affects AsiA interactions with RNAP and MotA (Yuan and Hochschild, 2009; Yuan et al., 2009). Hence it is still unclear what effect the frameshift mutation in ΦOT8s will have with regards to AsiA function.

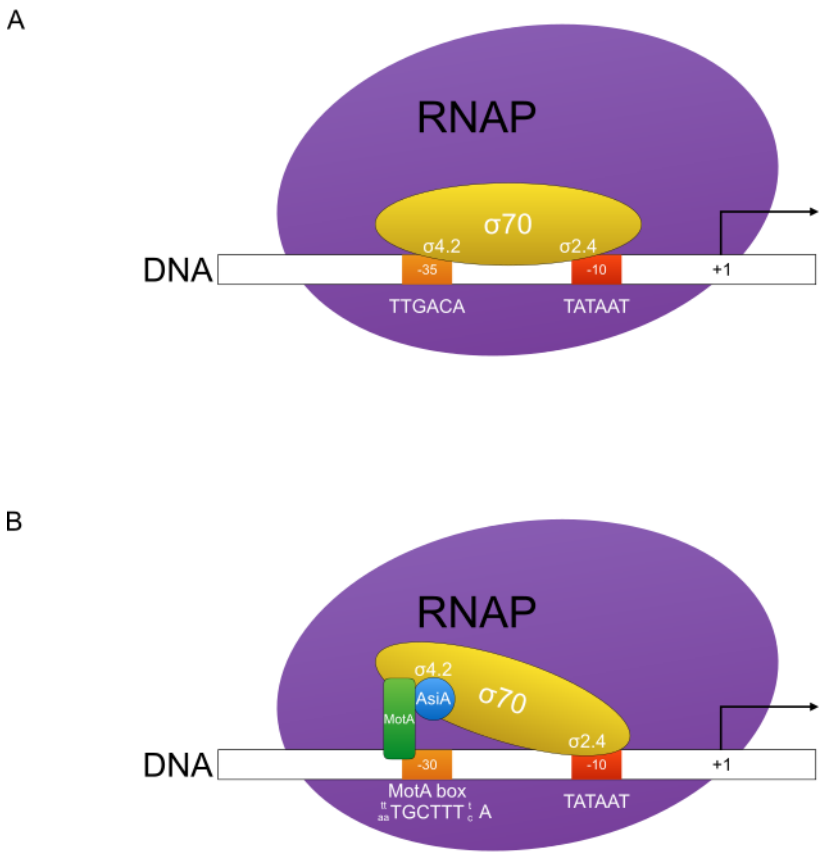



**Figure 6.7** σ appropriation in phage T4

**A**. In uninfected *E. coli*, the RNA polymerase holoenzyme recognises bacterial promoters through its interaction of the σ 70 subunit. **B**. During phage infection early proteins MotA and AsiA prevents this interaction. MotA interacts with a specific sequence called the MotA box present in the promoters of middle genes and activates the σ 70 subunit. AsiA binds to the 4.2 region and prevents it from recognising the -35 region sequences of both the early genes of T4 and host genes.

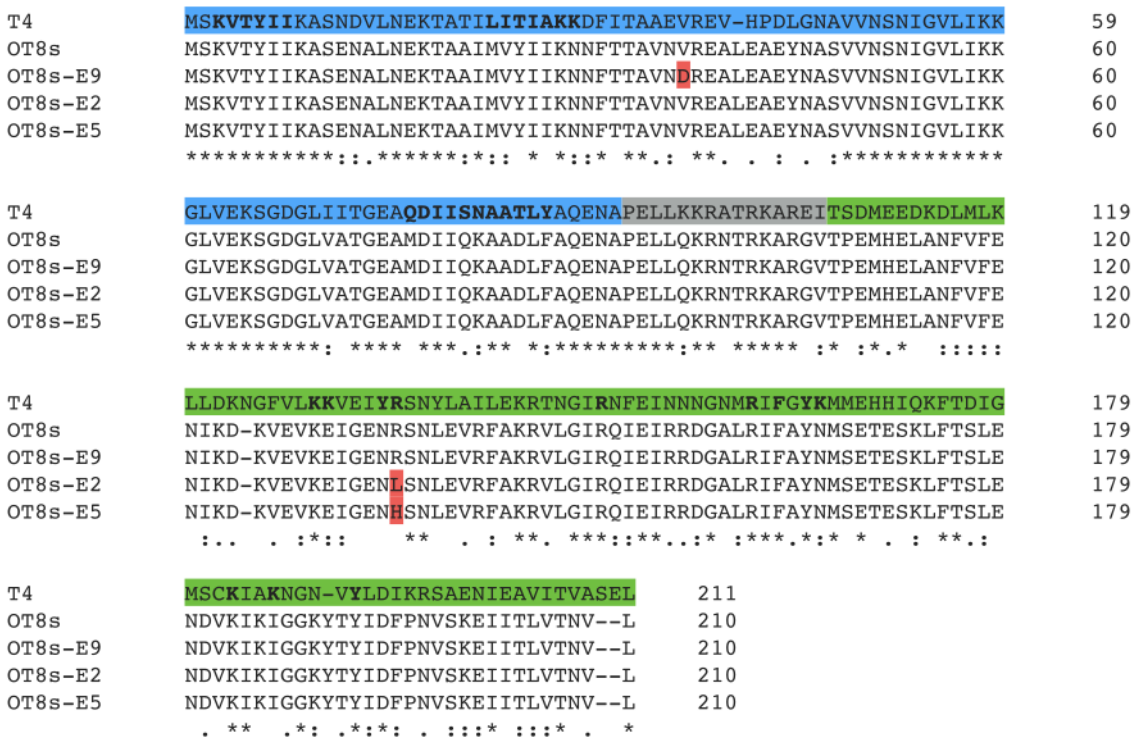



**Figure 6.8** MotA protein alignments of T4, ΦOT8s and ToxIN$_{Pa}$ escape mutants Residues in blue represent the N-terminal domain of T4 MotA. Those that are in bold in this domain are predicted to interact with the sigma 70 subunit. Residues in grey are a linker region connecting the N and C terminal domains. Residues in green represent the C-terminal domain area with those in bold representing putative DNA binding regions. Residues highlighted in red for the ΦOT8s escapes signify residues that were mutated.

```
T4          MNKNIDTVREIITVASILIKFSREDIVENRANFIAFLNEIGVTHEGRKLNQNSFRKIVSE   60
OT8s        MNKDIEIVREIITIASILIKFSREDIVEDRASFIGFLNEIGIKKDGRQLNQNSFRKLITN   60
OT8s-E1     MNKDIEIVREIITIASILIKFSREDIVEDRASFIGFLNEIGIKKRWSAAKSEFVQKTNYK   60
            ***:*: ******:**************:**.**.******:.:     :.: .:*   :

T4          LTQEDKKTLIDEFNEGFEGVYRYLEMYTNK 90
OT8s        LTAEEKKTLISEFNEGFENIYRHLAMYSNN 90
OT8s-E1     FNCRRKENT-D------------------- 70
            :. . *:.  .
```

**Figure 6.9** AsiA alignments of T4, ΦOT8s and escape mutant

Residues in red are differences from the wild type protein.

### 6.4.3 ΦOT8s is aborted and able to escape TenpIN$_{PI}$ in an EPI300 background

ΦOT8s was tested for sensitivity to the TenpIN$_{PI}$ system in both *Serratia* (pMUT13, pFR2) and DH5α (pFR2). The phage was aborted in both bacterial backgrounds but even at titres of over $10^{10}$ pfu.ml$^{-1}$ was unable to evolve escapes. However, when tested on EPI300 (pFR2) not only was ΦOT8s aborted but escapes plaques were observed (**Table 6.9**). To confirm that they were true escapes a single plaque was picked and plated on EPI300 (pFR2) and this bred true.

**Table 6.9** ΦOT8s can escape the TenpINPI system in EPI300

| Isolation host | Phage | EOPs | | |
|---|---|---|---|---|
| | | 39006 (pMUT13, pFR2) | DH5α (pFR2) | EPI300 (pFR2) |
| *Serratia* ATCC 39006 (pMUT13) | ΦOT8s (wt) | $<1x10^{-8}$ | $<1x10^{-8}$ | $3.7x10^{-5}$ |
| *E. coli* EPI300 (pFR2) | ΦOT8s-PI | ND | ND | 0.55 |

ND – not determined

## 6.6 Discussion

There is clear evidence that Type III TA system activation is at least partially dependent on the hosts in which they reside. Results from the coliphage experiments with DH5α and EPI300 show that, even in similar strains of *E. coli*, there can be differences in whether phages can escape the Type III TA system or not. Analysis of the genotypes of both DH5α and EPI300 do not show obvious mutations that would lead to the difference in phenotype with regards to Type III TA systems. One possible hypothesis is that even though both strains harbouring the same TA plasmid have different phenotypes it may be possible that the copy number is not the same between the two strains. The copy number of pBR322 is reported to be 15-20 copies per cell, which would be the expected number maintained in both DH5α and EPI300 (Lee et al., 2006). Plasmid copy number could ultimately affect the total number of TA complexes in the cell and has already been noted as a potential factor in Type III TA activation (Blower, 2009).

Although the reasons why certain coliphages could escape Type III TA systems in EPI300 but not in DH5α couldn't be defined in this work there are various conclusions that can be drawn from the results. One conclusion is that there are several mutations that can lead to phage escapes from Type III TA systems in one host that will not necessarily lead to escape in another. This may imply that the mutated phage products do not interact directly with the Type III TA systems but rather with a host factor that is linked to Type III TA activation.

One phage capable of escaping a Type III TA system in EPI300 is T6. This phage was aborted by the TenpIN$_{PI}$ system but was able to evolve mutations in the Alc transcriptional inhibitor that allow it to escape. In T4, Alc is involved in the takeover of the host transcriptional machinery by preventing transcription of non-modified cytosine bases. One obvious mechanism in which Alc might cause Type III TA activation is that, by generally repressing host transcription this will also prevent the transcription of the Type III antitoxin. The toxin of Type III TA systems is more stable than the cognate antitoxin counterpart, so generalised inhibition of transcription could eventually lead to a decrease in antitoxin levels thereby releasing the more stable toxin. It has also been shown that the Type III toxin-antitoxin full complex is not as common in vivo as

originally presumed but rather the toxin is constantly cleaving newly synthesised antitoxin (Short et al., 2018). Follow up experiments that measure the $TenpI_{Pl}$ levels during infection with T6 and an escape mutant should be performed to confirm whether this is true. If this were the case for $TenpIN_{Pl}$ activation in T6 in EPI300 it is curious that T6 doesn't evolve the same mutations when selected on DH5α as one would expect the same result. Clearly T6 interacts with host related factors differently between the two strains and so it is feasible that inhibition of the host transcriptional machinery will be different. It may be that while certain mutations could slightly impair the phage's ability to inhibit host transcription it may not be enough to escape in DH5α but is sufficient to allow the phage to escape in an EPI300 background.

The LamB dependent *Serratia* phage ΦOT8s provided the opportunity to study Type III TA system activation between two very different bacteria. The experiments with ΦOT8s showed that there is a similar mechanism of escape that allows these phages to escape $ToxIN_{Pa}$ in both *Serratia* and *E. coli* backgrounds. These escape phages were found to have mutations affecting two proteins, MutA and AsiA, which function together in a process called sigma appropriation. Similar to the *alc* mutations found in T6 escapes from the $TenpIN_{Pl}$ system, MotA and AsiA may work in a similar process because by stopping transcription of host RNA, antitoxin levels would drop leading to an active toxin. Thus by disrupting MotA and AsiA function, host transcription, which would include the antitoxin is not as strongly affected thus Abi is avoided. An alternative mode of Type III TA system activation in this case is that, as MotA and AsiA promote the transcription of T4 middle gene products, that one of these middle gene products could be involved in activation of the system. The *asiA* mutation is also of interest as this is not the first time this protein has been investigated with regards to $ToxIN_{Pa}$ system escape. In fact, other *Serratia Tevenvirinae* phages, ΦCHI14, ΦXF20 and ΦCBH8 have been isolated previously in this laboratory that are able to escape $ToxIN_{Pa}$ through mutations in *asiA* (Chen et al., 2017). Other mutations identified in the latter phages that can allow escape of $ToxIN_{Pa}$ include a large deletion encompassing several tRNA and genes of unknown functions, as well as mutations in a single gene, *orf84* encoding a protein of unknown function (Chen et al., 2017). ACT alignments of T4 and these *Serratia* phages showed no homology within the large deletion locus making it difficult to suggest its function. Furthermore, the proteins in that region in T4 all have unknown functions although work on characterising these proteins in T4 has

shown that several of these proteins are in fact toxic (Miller et al., 2003). This suggests that these phage products must interact with host factors in some important way. It is likely that overexpression of some of these *Serratia* phage products is toxic, indeed this is an emerging theme in several studies of Type III TA systems (Blower et al., 2017). The M1-23 protein, Alc, MotA and AsiA are all highly toxic proteins when expressed in bacteria (Blower et al., 2017). ΦOT8s does not have a homolog of *orf84*, however it has homologs of four of the proteins deleted in the large deletion region of escapes. Whether these four proteins individually or collectively are able to activate $ToxIN_{Pa}$ has yet to be determined.

One interesting point to note is that ΦCHI14, ΦX20 and ΦCBH8 are aborted by the $TenpIN_{Pl}$ system and unable to escape (Chen et al., 2017). ΦOT8s was also shown to be aborted by the $TenpIN_{Pl}$ system and unable to escape in a *Serratia* 39006 or DH5α background, however it was able to escape in an EPI300 background. It would be interesting to see if transfer of the *Serratia* phage receptor allows infection of these phages in an EPI300 background and whether these phages are then aborted and able to escape. Another *Tevenvirinae* phage that is able to escape the $TenpIN_{Pl}$ system in an EPI300 background is T6 which, as discussed previous, has mutations in the *alc* gene. There are several T6 escapes where the mutation is not in *alc* and are not identified. Further work is required to determine the nature of these mutations which could of course, be *motA* or *asiA* mutations. Furthermore, there are now a bank of coliphages that have been identified that are able to evolve escapes from Type III TA systems in an EPI300 background which will be an invaluable source for further study.

# Chapter 7 Isolation of a broad host range superfamily of phages capable of infecting insect, animal, human and plant pathogens

## 7.1 Introduction

The Type III TA system, TenpIN$_{Pl}$ has been shown to be an effective abortive infection system based on previous work and in this study. This TA system originated from the Gram negative insect pathogen, *Ph. luminescens* TT01 where it is located chromosomally (Blower et al., 2012b). Expression of this TA system in plasmids showed it was capable of aborting phages isolated from *P. atrosepticum*, *E. coli* and *Serratia* (Blower et al., 2012b, 2017; Chen et al., 2017). However, there are no phages isolated on *Ph. luminescens* itself that are aborted by the TenpIN$_{Pl}$ system (Rao, 2014). Previous work from members of this lab has created a TenpIN$_{Pl}$ deletion strain of TT01, FRA2 (Rao, 2014). Phages isolated on this strain were tested on the wild type TT01 strain (containing a single chromosomal copy of the TenpIN$_{Pl}$ system) as well as TT01 transformed with pFR2. None of the phages isolated were aborted which was not expected as TenpIN$_{Pl}$ had been shown to abort a wide range of phages from different bacterial backgrounds. These results suggest that perhaps the normal physiological role of the TenpIN$_{Pl}$ system is not to function as an abortive infection system in *Ph. luminescens* but that it functions as an abortive infection system when in multiple copies in other bacteria.

There has only been one phage published as infecting *P. luminescens*. However, there is limited data on this phage and no genomic sequence has been published (Poinar et al., 1989). *Ph. luminescens* is referred to in this publication as *Xenorhabdus luminescens* which was what it was previously called (Boemare et al., 1993). Sequencing of the TT01 genome has revealed numerous prophage elements, however, none have been reported to be produce an intact phage particle(Gaudriault et al., 2004). Thus there are no published genomes of TT01 phages to date. This chapter describes work on the isolation of novel TT01 phages and their characterisation with respect to Abi and other aspects.

## 7.2 All TT01 phages produce faint turbid plaques and are insensitive to both the ToxIN$_{Pa}$ and TenpIN$_{Pl}$ system

From 2012 to 2016 several members of the lab had isolated phages from either the Milton sewage works or the river Cam using the *Ph. luminescens* strain FRA2. This formed a library of approximately 80 phages of which 31 were isolated during this study, ΦRAY0-30. All phages from this library produced extremely faint, turbid plaques on TT01, a phenotype seen previously (Rao, 2014). Furthermore, ΦRAY0-30 were not aborted by wild type TT01 or TT01 carrying pTA46 or pFR2.

## 7.3 TT01 phages are motility dependent *Chiviruses*

As there were no data published on TT01 phages and although none showed any sensitivity to the two Type III TA systems, eight phages from the TT01 library were chosen to be genomically sequenced. The phages were selected so that the dates on which they were isolated were as far apart as possible, in the hope of selecting the most diverse phages as possible. Sequencing results for the eight TT01 phages revealed that they all were similar phages that belonged to the phage genus *Chivirus* (**Table 7.1**). The TT01 phages had linear genomes that were a little under 60,000 base pairs long encoding 73 to 75 predicted ORFs and shared over 89.9% identity homology with each other (**Table 7.1, 7.2**). Their genomes also closely match that of the founding member phage Chi, a *Siphoviridae* member that was isolated on a strain of *Salmonella* (Sertic and Boulgakov, 1936). One of the sequenced phages, ΦRAY26 was viewed by TEM and was also shown to be a member of the *Siphoviridae*, characterised by a long non-contractile tail. (**Figure 7.1**). During the period of this study a member from the lab of Gordon Dougan at the Wellcome Trust Sanger Institute, Derick Pickard, isolated another *Chivirus* on *S.* Typhi (unpublished). The annotated genome sequence of this phage, YSD1 was provided by Derick as well as a phage lysate which has been used in this study for comparisons with the TT01 *Chiviruses* (**Table 7.1, 7.2**).

Genomic analysis of the TT01 phages showed that their genomes could be spilt into six different regions based on the functions of the protein encoded within those regions (**Figure 7.2**). Alignments of the genomes of the TT01 phages showed that the regions that code for structural proteins were highly conserved with the greatest variability arising in the region associated with DNA modification and recombination (**Figure 7.3**).

The genomes of *Chiviruses* deposited into Genbank showed that these phages are commonly isolated worldwide with all phages isolated on various strains of *Salmonella*, bar one, which was isolated on *E. coli* (**Table 7.3**). Sequence alignments of these genomes with a representative member (ΦRAY26) also showed that these phages share similar genomic identities (**Table 7.3**). Alignments made with the TT01 phages and these *Chiviruses* also show a similar pattern where variability between phages is in the region associated with DNA modification and recombination (**Figure 7.4**).

Phage Chi was previously shown to be dependent on motile flagella (Samuel et al., 1999) thus the library of TT01 phages was tested on two motility mutants of TT01, Δ*flgG* and Δ*motAB* (Easom and Clarke, 2008). As *flgG* encodes a structural component of the flagellar machinery deletion of this gene leads to the complete absence of flagella. On the other hand, Δ*motAB* encodes part of the motor component of the flagella so deletion of this gene still leads to the production of flagella but they are unable to rotate and thus the strain is non motile. Phages from the TT01 library were tested on these motility mutants but were found to be unable to infect either of the strains. An adsorption assay performed with ΦRAY26 showed that there was no adsorption to either mutants (**Figure 7.5**). This meant that, in the case of the Δ*motAB* mutant, even though there were flagella present the phage could not adsorb onto it unless it was rotating. The reasons for the poor plaque formation of these phages on TT01 was investigated. One explanation is that the flagella are differentially expressed. To test this a 20 µl spot of neat TT01 phage, ΦRAY26 was spotted onto the centre of a lawn of TT01. The plates were incubated and a turbid spot appeared. The bacteria from this spot, which had been exposed to phage were streaked to single colonies and were tested for resistance to the phage. These bacteria were still sensitive to the phage suggesting that they might have gained temporary resistance to the phage by halting motility but, in the absence of phage, motility could be switched on again.

As no TT01 phage could infect any of the motility mutants, these mutants were used in new enrichment assays to try to isolate novel TT01 phages that are not dependent on motility. Interestingly, with over 10 different water samples tested it was not possible to isolate any TT01 phage on either of the motility mutants despite facile isolation of flagellum-dependent phages on TT01. This implies that TT01 motility independent

phages are extremely rare. TT01 itself is an insect pathogen found in soil so it may be unlikely to find high populations of these bacteria in water sources. However, TT01 *Chiviruses* are relatively easy to isolate,essentially every sample of water collected that had been enriched contains these viruses. Therefore, one hypothesis is that these flagellum-dependent *Photorhabdus* phages replicate on alternative hosts in the natural environment.

**Table 7.1** TT01 phages share high genome sequence homology to each other, to phage Chi and to YSD1

| | Chi | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | YSD1 |
|---|---|---|---|---|---|---|---|---|---|---|
| Chi | 100 | 93.5 | 94.9 | 91.9 | 95.2 | 94.7 | 92.6 | 89.9 | 96.3 | 94.3 |
| ΦLT | 93.5 | 100 | 95.1 | 90.8 | 94.8 | 90.8 | 91.3 | 90 | 95.2 | 93.1 |
| ΦRAY0 | 94.9 | 95.1 | 100 | 92.1 | 95 | 92.9 | 91.7 | 91.8 | 95 | 94.1 |
| ΦRAY1 | 91.9 | 90.8 | 92.1 | 100 | 92.8 | 91.9 | 90.7 | 92.9 | 90.7 | 91.1 |
| ΦRAY19 | 95.2 | 94.8 | 95 | 92.8 | 100 | 92.3 | 91.7 | 90.4 | 95.1 | 94.3 |
| ΦRAY26 | 94.7 | 90.8 | 92.9 | 91.9 | 92.3 | 100 | 91.4 | 91.3 | 93.3 | 92.9 |
| ΦRAY27 | 92.6 | 91.3 | 91.7 | 90.7 | 91.7 | 91.4 | 100 | 90.7 | 93.3 | 90.6 |
| ΦJB2 | 89.9 | 90 | 91.8 | 92.9 | 90.4 | 91.3 | 90.7 | 100 | 90 | 92.2 |
| ΦJB15 | 96.3 | 95.2 | 95 | 90.7 | 95.1 | 93.3 | 93.3 | 90 | 100 | 93.7 |
| YSD1 | 94.3 | 93.1 | 94.1 | 91.1 | 94.3 | 92.9 | 90.6 | 92.2 | 93.7 | 100 |

The colour for each box has been generated by conditionally formatting using Microsoft Excel. The more “green” a box is the higher genomic sequence identity it shares with the corresponding genome. On the other hand, the more “red” a box is, the less the genomic sequence identity it shares with the corresponding genome.

**Table 7.2** Genomic data of TT01 phages and *Salmonella* phages Chi and YSD1

| Phage | Isolation date | Genome length (bp) | Predicted ORFs |
|---|---|---|---|
| Chi | 1936 | 59,578 | 75 |
| ΦLT | 2012 | 59,911 | 74 |
| ΦRAY0 | 5th Feb 2015 | 59,163 | 74 |
| ΦRAY1 | 13th Apr 2015 | 59,136 | 75 |
| ΦRAY19 | 13th May 2015 | 59.147 | 73 |
| ΦRAY26 | 30th Jul 2015 | 59,486 | 74 |
| ΦRAY27 | 11th Aug 2015 | 59,906 | 75 |
| ΦJB2 | Apr 2016 | 58,922 | 74 |
| ΦJB15 | Apr 2016 | 59.824 | 74 |
| YSD1 | 2016 | 58,916 | 71 |

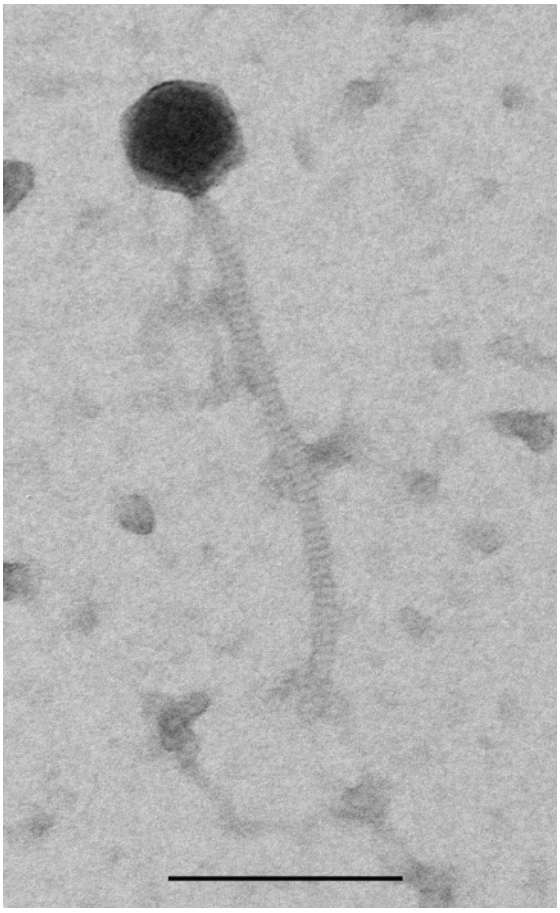


**Figure 7.1** TEM image of ΦRAY26

The long non contractile tail observed from the image confirms that ΦRAY26 is a *Siphoviridae* member concurring with the genomic data obtained. The black bar represents 100 nm.

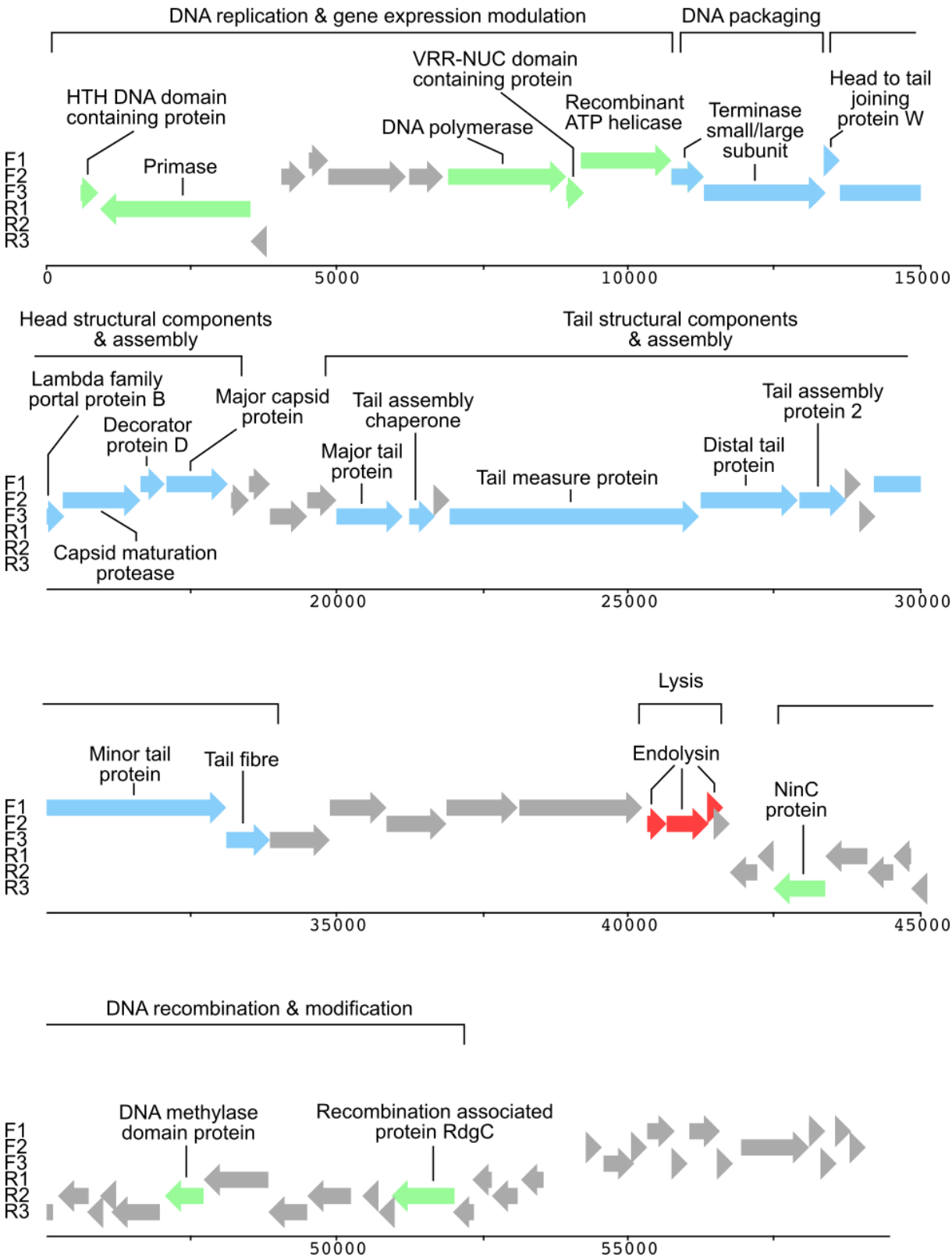



**Figure legend on next page**

**Figure 7.2** Genomic map of ΦRAY26

The genome of ΦRAY26 and other *Chiviruses* can be split into six regions based on the functions of the proteins in the region. Arrows in green, blue and red represents ORFs involved in metabolism, virion structure and lysis respectively. Arrows in grey depict ORFs with unknown functions. F1-3 and R1-3, refer to the six different reading frames.

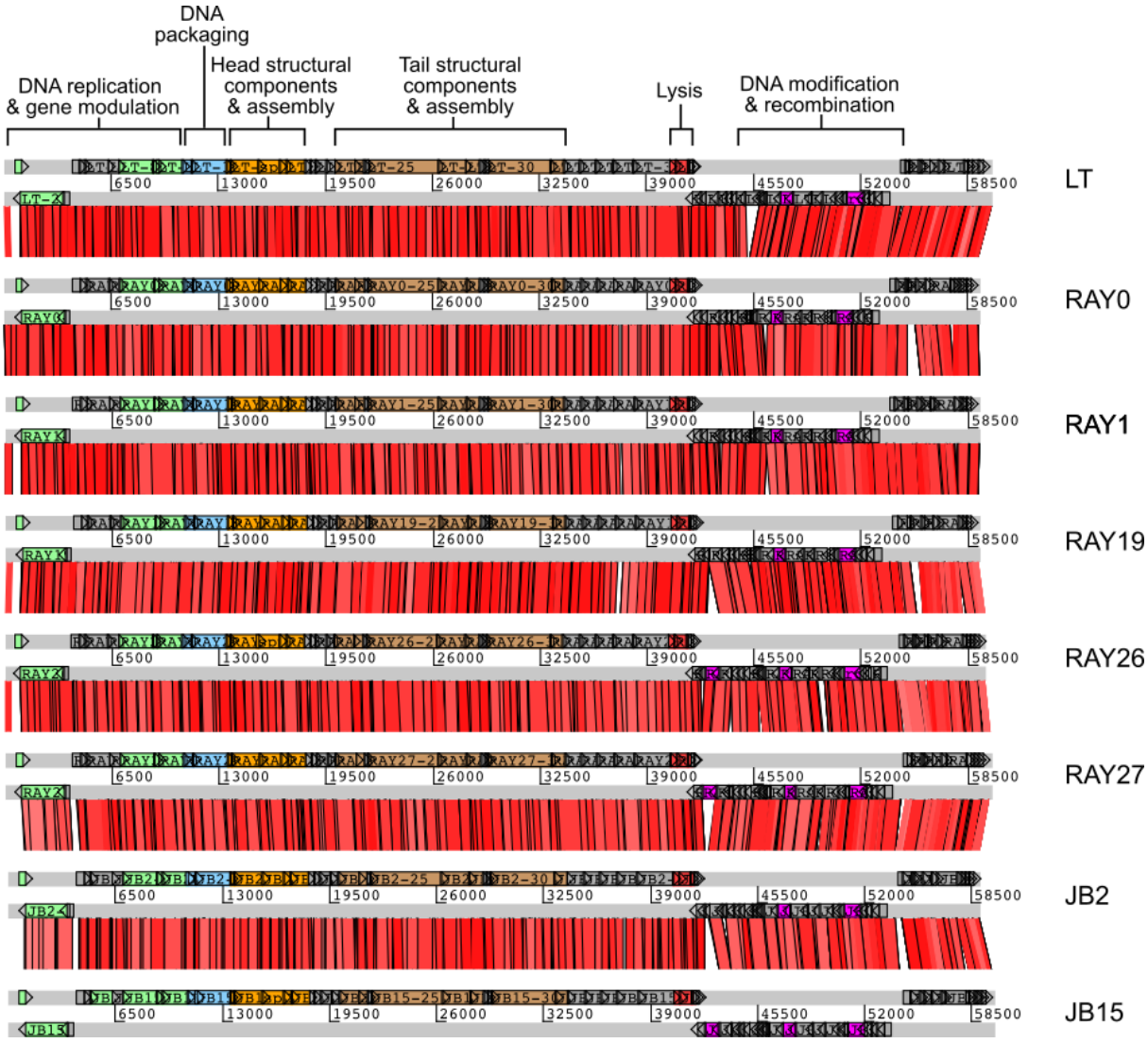



**Figure 7.3** Genome alignments of TT01 phages

The variability between phages tend to be in the region encoding ORFs involved in DNA modification and recombination. The ACT comparison file was generated using tblastx with a cut off value of $1x10^{-4}$. Red bars show areas of high amino acid similarity.

**Table 7.3** Complete *Chivirus* genomes deposited in Genbank

| Phage | Host isolated on | Size | Year isolated/genome submitted | Location/Lab location | % identity to ΦRAY26 |
|---|---|---|---|---|---|
| Chi | *S. enterica* | 59,578 | 1936 | Paris, France | 94.7 |
| Φ37 | *S. gallinarum* | 60,216 | 2002 | Pune, India | 90.5 |
| Φ35 | *S. gallinarum* | 55,391 | 2002 | Pune, India | 86.4 |
| BP12C | *S. enterica* Hadar | 60,606 | 2003 | Montreal, Canada | 90.8 |
| SPN19 | *Salmonella* spp. | 59,203 | 2011 | Seoul, South Korea | 90.3 |
| iEPS5 | *S*. Typhimurium SL1344 | 59,254 | 2013 | Seoul, South Korea | 92.0 |
| FSL SP-124 | *Salmonella* spp. | 59,245 | 2013 | New York, USA | 91.4 |
| FSL SP-088 | *Salmonella* spp. | 59,454 | 2013 | New York, USA | 91.9 |
| FSL SP-030 | *Salmonella* spp. | 59,746 | 2013 | New York, USA | 88.9 |
| FSL SP-039 | *Salmonella* spp. | 59,815 | 2013 | New York, USA | 88.8 |
| 118970 Sal1 | *Salmonella enterica* | 59,518 | 2014 | Portici, Italy | 91.4 |
| Utah | *E. coli* SKB178 | 59,024 | 2016 | Texas, USA | 92.7 |

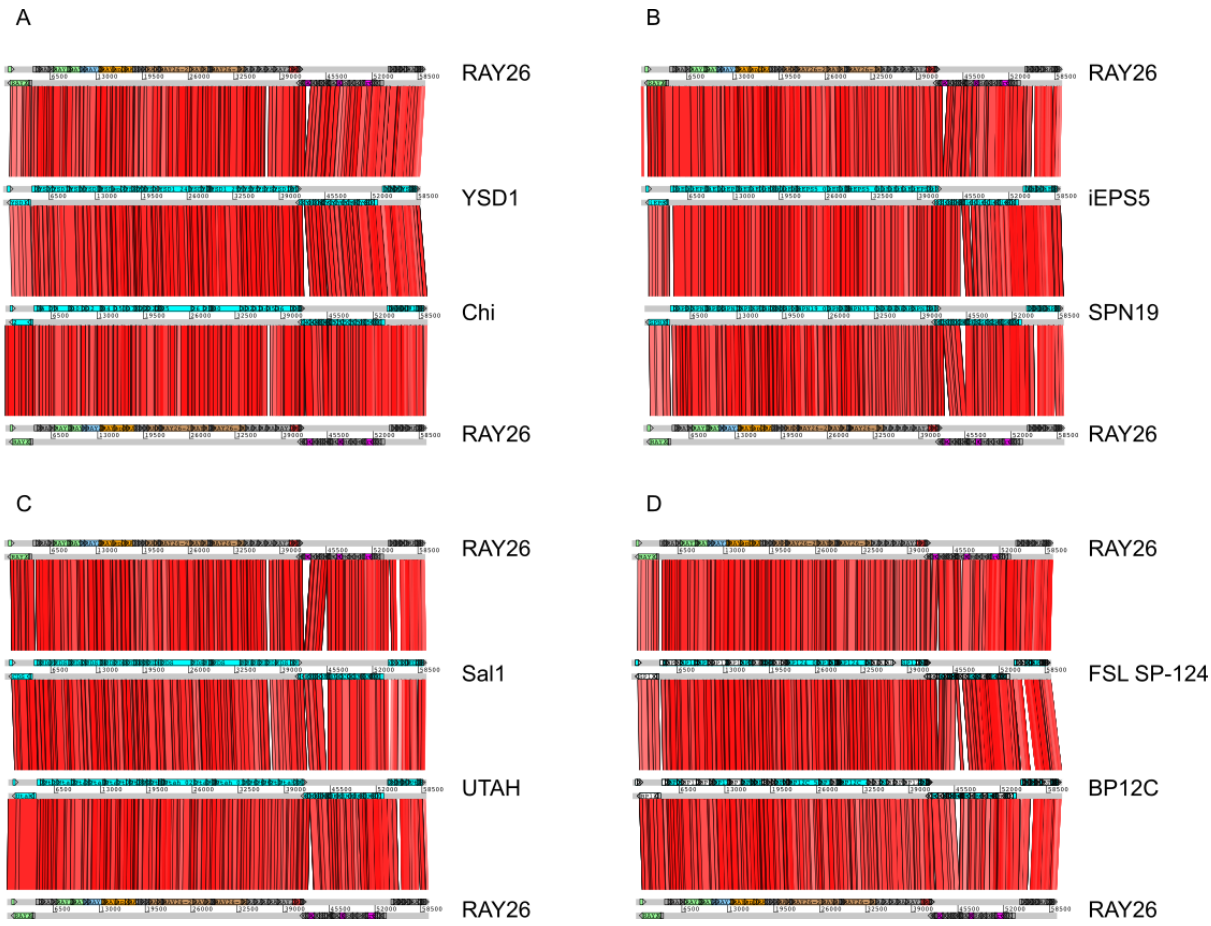



**Figure 7.4** *Chiviruses* are very well conserved apart from the regions involved in DNA modification and recombination

**A**. ΦRAY26 compared with the oldest and founding member of the *Chiviruses*, Chi and YSD1. **B**. ΦRAY26 compared with iEPS5 and SPN19. **C**. ΦRAY26 compared with Sal1 and the only *Chivirus* isolated on *E. coli*, phage Utah. **D**. ΦRAY26 compared with FSL SP-124 and BP12C. The ACT comparison file was generated using tblastx with a cut off value of $1x10^{-4}$. Red bars show areas of high amino acid similarity.

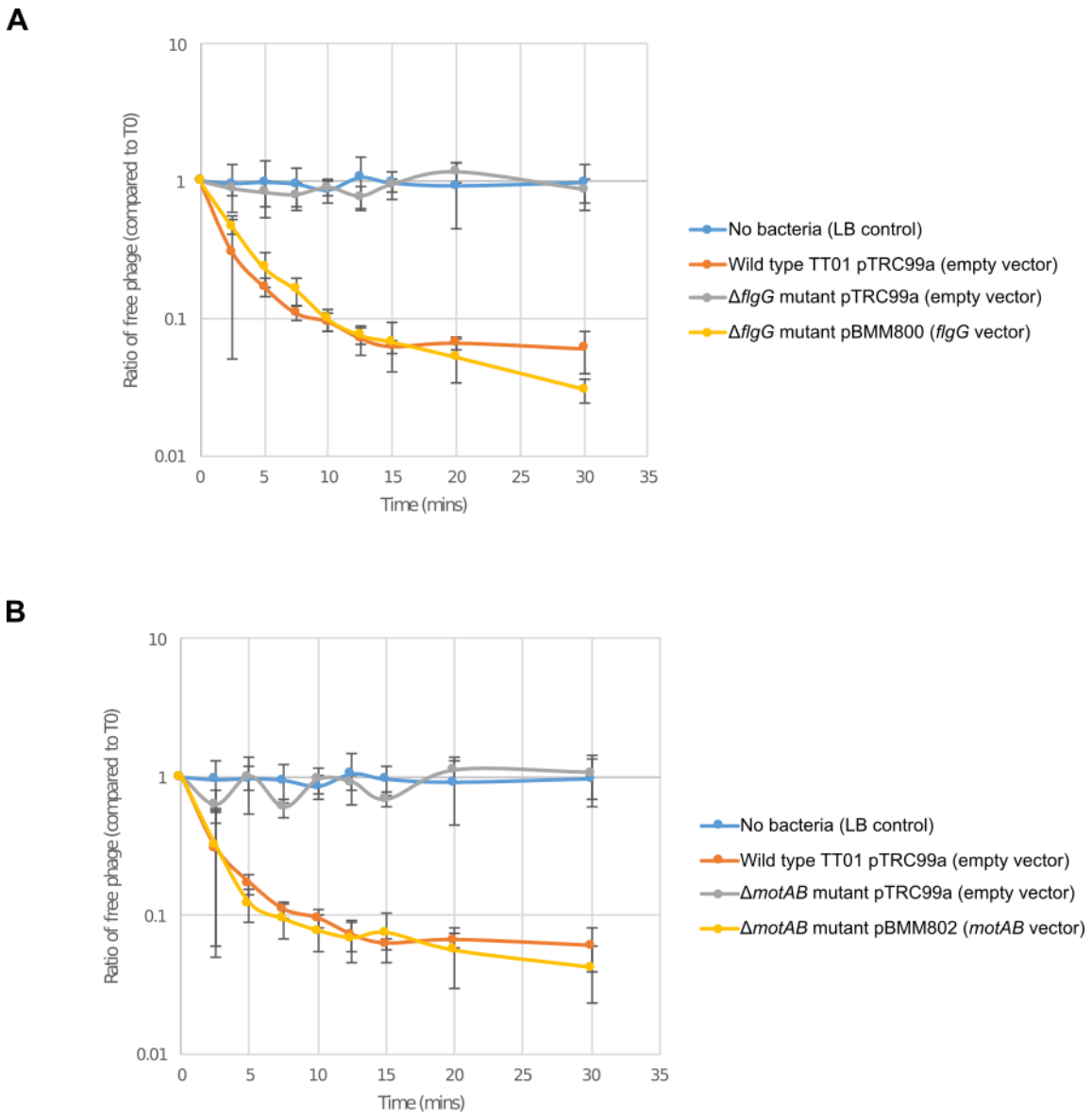



**Figure 7.5** ΦRAY26 does not adsorb to TT01 motility mutants

**A**. Adsorption assay of ΦRAY26 and the Δ*flgG* mutant. **B**. Adsorption assay of ΦRAY26 and the Δ*motAB* mutant. Error bars represent the standard deviation with N=3. A MOI of 0.01 was used in both assays.

### 7.4 TT01 exhibit host range promiscuity but in a phage-dependent fashion

The TT01 phages all shared high homology to the *Chiviruses* which have been mostly isolated on various *Salmonella* strains bar one which was isolated on an *E. coli* strain. The YSD1 phage that was isolated on a *Salmonella* strain was tested on TT01 and was able to infect it. Likewise, TT01 phages that were tested on *S.* Typhimurium strain 5383 were found to be able to infect it (**Table 7.4**). Interestingly, not all of the TT01 phages infected 5383 equally efficiently. While most phages infect *Salmonella* 5383 just as efficiently as TT01, one phage ΦRAY0 infected 5383 poorly (**Table 7.4**). Plaques however, could still be picked and plating this phage mutant, ΦRAY0-5383, showed that it could then replicate just as efficiently on TT01 and 5383. To confirm that this result was not caused by restriction and modification (RM), a plaque of ΦRAY0-5383 grown on TT01 was picked and plated on TT01 and 5383 again. The phage from this new plaque, ΦRAY0-5383-TT01, showed a three log drop in titre on 5383 compared to TT01, the previous host on which it replicated, suggesting that it is likely not the result of RM (**Table 7.5**). Additionally, for all host range assays, a plaque of YSD1 grown on TT01 was used because the *Salmonella* strain that was host for YSD1 was not available.

Although isolated on *Salmonella*, phage Chi was also able to infect motile strains of *E. coli* and *Serratia* (Iino and Mitani, 1967; Schade and Adler, 1967). The TT01 phages were also shown to be able to infect motile strains of *E. coli* and *Serratia* (**Table 7.4**). Three *E. coli* strains were tested (DH5α, W3110 and MG1655) where only MG1655 is motile; the TT01 phages could only infect MG1655. Two motile strains of *Serratia* were also used, DB10 and ATCC 39006, however the TT01 phages could only infect DB10 suggesting that there may still be specificity involved in adsorption to the different flagella. Furthermore, only one phage, (ΦRAY1), plated efficiently on DB10. All other TT01 phages, as well as YSD1, replicated poorly in this host. However, picking a plaque of ΦRAY26 grown on DB10 (ΦRAY26-DB10) and plating it back on DB10 and TT01, showed that it could then plate just as efficiently on both hosts (**Table 7.5**). Several other strains of bacteria available in the lab were also tested for their susceptibility to the TT01 phages. This included several Gram positive bacteria, *S. aureus* and three *B. thuringiensis* strains, and Gram negatives, *Citrobacter rodentium*,

*Chromobacter violaceum*, *Pseudomonas aeruginosa* and *Vibrio harveyi* but none of the TT01 phages were able to form plaques on these hosts (**Table 7.4**).

**Table 7.4**. Host range of Chi-like phages against misc. bacteria

| Bacteria | Strain | Chi-like phages EOPs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | ΦYSD1 |
| *P. luminescens laumondii* | TT01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| *S.* Typhimurium | 5383 | 6.43 | $1.13x10^{-7}$ | 0.11 | 1.72 | 4.40 | 2.50 | 0.02 | 4.95 | 2.46 |
| *E. coli* | MG1655 | 8.00 | 2.64 | 0.84 | 2.02 | 3.71 | 2.04 | 0.02 | 8.51 | 3.74 |
| *E. coli* | W3110 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *E. coli* | DH5α | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *C. rodentium* | ICC169 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *C. violaceum* | CV026 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *S. marcescens* | DB10 | $7.74x10^{-6}$ | $4.36x10^{-8}$ | 2.26 | $9.34x10^{-7}$ | $2.76x10^{-6}$ | $7.42x10^{-8}$ | $2.31x10^{-7}$ | $6.26x10^{-6}$ | $8.91x10^{-7}$ |
| *Serratia* | ATCC 39006 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *P. aeruginosa* | PA01 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *V. harveyi* | BB120 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *S. aureus* | H | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *B. thuringiensis* | HD-1 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *B. thuringiensis* | HD-73 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |
| *B. thuringiensis* | JH642 | $<2.62x10^{-8}$ | $<1.36x10^{-8}$ | $<2.18x10^{-8}$ | $<1.44x10^{-8}$ | $<4.54x10^{-8}$ | $<4.76x10^{-9}$ | $<7.70x10^{-9}$ | $<2.80x10^{-8}$ | $<6.23x10^{-8}$ |

**Table 7.5** Putative host range mutants of selected TT01 phages

| Phage | Host | EOP |
|---|---|---|
| ΦRAY0 | TT01 | 1 |
| | 5383 | $1.13x10^{-7}$ |
| ΦRAY-5383 | TT01 | 0.79 |
| | 5383 | 1 |
| ΦRAY-5383-TT01 | TT01 | 1 |
| | 5383 | $2.44x10^{-3}$ |
| ΦRAY26 | TT01 | 1 |
| | DB10 | $2.76x10^{-6}$ |
| ΦRAY-DB10 | TT01 | 0.49 |
| | DB10 | 1 |

TT01 is an insect pathogen, so to see if the TT01 phages could also infect other insect pathogens, they were tested on a small library of available insect pathogens (**Table 7.6**). Results from these experiments confirmed that many of the insect pathogens tested were indeed vulnerable to TT01 phage infection. Again, a pattern emerged where particular TT01 phages could infect certain strains more efficiently than other TT01 phages.

Several other libraries of bacteria were also tested for their susceptibility to TT01 phage infection. One of these libraries was a collection of phytopathogens which included various *P. atrosepticum*, *P. carotovorum* and *Dickeya* strains. A selection of these strains was tested but none were susceptible to TT01 phage infection (**Table 7.7**). The TT01 phages had previously been shown to infect *S. marcescens* strain Db10, so a library of other *S. marcescens* strains were tested (**Table 7.8**). Several of these strains were susceptible to TT01 phage infection and yet again there were differences in efficiency of infection between phages on certain strains. The host ranges of these phages also appears to be random with no obvious grouping. A final library was tested consisting of bacteria isolated from the rhizosphere that had been classified by FAME (Berg et al., 2002). Many of the bacteria were susceptible to TT01 phage infection suggesting that there are many bacterial habitats that harboured bacteria sensitive to the TT01 phages (**Table 7.9**).

**Table 7.6** Host range of Chi-like phages against insect pathogen library

| Bacteria | Strain | Chi-like phages EOPs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | ΦYSD1 |
| *P. luminescens laumondii* | TT01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| *P. luminescens akhurstii* | LN2 | 1.70 | 0.40 | 1.15 | 0.29 | 0.85 | 0.38 | 0.046 | 1.25 | 0.39 |
| *P. luminescens laumondi* | HP88 | 0.08 | 0.23 | 0.34 | 0.22 | 0.28 | 0.23 | 0.48 | 0.20 | 0.24 |
| *P. luminescens akhurstii* | P2H | $4.76\times10^{-6}$ | $1.36\times10^{-8}$ | 0.01 | $9.25\times10^{-6}$ | $5.46\times10^{-4}$ | 0.07 | 0.02 | 5.36 | $4.10\times10^{-6}$ |
| *P. luminescens akhurstii* | K4D | 1.82 | 0.33 | 1.18 | 0.21 | 0.57 | 0.52 | 1.50 | 2.75 | 0.62 |
| *P. asymbiotica* | ATCC 43950T | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *P. asymbiotica* | Kingcliffe | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *X. feltiae* | - | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *X. nematophila* | - | $9.88\times10^{-5}$ | $7.70\times10^{-7}$ | $4.33\times10^{-6}$ | $3.91\times10^{-6}$ | $8.41\times10^{-8}$ | 0.43 | 1.55 | 0.67 | $3.16\times10^{-7}$ |
| *Photorhabdus* | Mexico | 2.07 | 0.79 | 0.93 | 0.30 | 1.06 | 0.44 | 1.38 | 2.08 | 0.99 |
| *Photorhabdus* | Egypt | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |

**Table 7.7** Host range of Chi-like phages against phytopathogenic bacteria

| Bacteria | Strain | Chi-like phages EOPs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | ΦYSD1 |
| *P. luminescens laumondii* | TT01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| *P. atrosepticum* | SCRI1043 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *P. atrosepticum* | SCRI1039 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *P. carotovora* | 39048 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *D. solani* | MK10 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *D. zeae* | NCPPB 3532 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *D. chrysanthemi* | NCPPB 402 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *D. dieffenbachiae* | NCPPB 2976 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *D. dianthicola* | NCPPB 453 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *D. paradisiaca* | NCPPB 2511 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |

**Table 7.8** Host range of Chi-like phages against *Serratia* library

| Bacteria | Strain | Chi-like phages EOPs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | ΦYSD1 |
| *P. luminescens laumondii* | TT01 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *S. marcescens* | 1695 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 0006 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 1047 | 0.52 | 1.22 | 2.45 | $3.58\times10^{-4}$ | 1.55 | 0.55 | 2.54 | 1.83 | 0.27 |
| *S. marcescens* | 0026 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 4028 | $2.00\times10^{-4}$ | $2.55\times10^{-3}$ | $7.57\times10^{-7}$ | $2.11\times10^{-6}$ | $3.20\times10^{-7}$ | 1.31 | 1.52 | 1.08 | 0.31 |
| *S. marcescens* | 3078 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 0038 | 0.07 | 0.04 | $6.73\times10^{-4}$ | $5.42\times10^{-6}$ | 0.02 | 0.02 | 0.06 | 0.04 | 0.30 |
| *S. marcescens* | 1024 | 6.17 | 1.61 | 5.54 | $1.73\times10^{-4}$ | 2.02 | 2.06 | 2.37 | 5.05 | 2.56 |
| *S. marcescens* | 3127V | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 0035 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $1.53\times10^{-7}$ | $7.46\times10^{-7}$ | $<2.80\times10^{-8}$ | $1.46\times10^{-7}$ |
| *S. marcescens* | 3888 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 0028 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 3255 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 0040 | $<2.62\times10^{-8}$ | $4.55\times10^{-7}$ | $4.93\times10^{-5}$ | $<1.44\times10^{-8}$ | $1.73\times10^{-7}$ | $9.82\times10^{-6}$ | $2.56\times10^{-6}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 2595 | $<2.62\times10^{-8}$ | 0.08 | 0.07 | $1.49\times10^{-6}$ | 0.07 | $7.69\times10^{-9}$ | 1.64 | $2.90\times10^{-3}$ | 0.01 |

**Table 7.8** continued. Host range of Chi-like phages against *Serratia* library

| Bacteria | Strain | Chi-like phages EOPs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | ΦYSD1 |
| *P. luminescens laumondii* | TT01 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| *S. marcescens* | 0030 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 3127 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | 3078V | $8.36\times10^{-4}$ | $4.91\times10^{-3}$ | $5.44\times10^{-4}$ | $5.22\times10^{-7}$ | $5.48\times10^{-3}$ | $4.77\times10^{-3}$ | $9.37\times10^{-3}$ | $1.03\times10^{-3}$ | $1.54\times10^{-4}$ |
| *S. marcescens* | ATCC274 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | S6 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *S. marcescens* | SM365 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |

**Table 7.9** Host range of Chi-like phages against rhizosphere isolates library

| Bacteria (determined by FAME) | Strain | Chi-like phages EOPs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | ΦYSD1 |
| *P. luminescens laumondii* | TT01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| *Proteus vulgaris* | 10Ep11 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $8.85\times10^{-8}$ | $<1.44\times10^{-8}$ | 0.02 | $2.16\times10^{-3}$ | 0.02 | $<2.80\times10^{-8}$ | $1.77\times10^{-3}$ |
| *Serratia grimesii* | 9Ez25 | $2.81\times10^{-4}$ | $5.81\times10^{-8}$ | 0.21 | 0.15 | 1.09 | 1.40 | 1.73 | $1.69\times10^{-7}$ | $8.70\times10^{-8}$ |
| *Serratia grimesii/Pantoea agglomerans* | 9Ez29 | $2.36\times10^{-4}$ | 8.17x10-8 | 0.13 | $5.15\times10^{-8}$ | $2.79\times10^{-7}$ | 1.51 | 2.32 | $9.95\times10^{-8}$ | $6.23\times10^{-8}$ |
| *Xenorhabdus nematophila* | 9Ec15 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Kluyvera cryocrescens* | 2Kr27 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Enterobacter intermedius/Serratia grimesii* | 4Rc14 | $<2.62\times10^{-8}$ | $6.44\times10^{-8}$ | 0.11 | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | 0.99 | 0.02 | $<6.23\times10^{-8}$ |
| *Enterobacter intermedius* | 4Rz1 | $2.11\times10^{-6}$ | $6.67\times10^{-9}$ | 0.04 | $5.52\times10^{-8}$ | $<4.54\times10^{-8}$ | $8.54\times10^{-4}$ | 1.23 | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Pantoea agglomerans* | 4RZ11 | $6.43\times10^{-8}$ | $3.94\times10^{-8}$ | 0.17 | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | 0.05 | 0.79 | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Pantoea agglomerans/Serratia grimesii* | 9Rz4 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Pantoea agglomerans/Klebsiella pnuemoniae* | 3Rc14 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Proteus vulgaris/Serratia grimesii/Pantoea agglormerans* | 4Rr2 | $7.19\times10^{-7}$ | $1.78\times10^{-8}$ | 0.08 | $<1.44\times10^{-8}$ | $6.31\times10^{-5}$ | $3.80\times10^{-6}$ | 0.61 | $4.66\times10^{-8}$ | $5.40\times10^{-8}$ |

**Table 7.9** continued. Host range of Chi-like phages against rhizosphere isolates library

| Bacteria (determined by FAME) | Strain | Chi-like phages EOPs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | ΦYSD1 |
| *P. luminescens laumondii* | TT01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| *Proteus vulgaris* | 4Rr2 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia fonticola* | 3Rc3 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia grimesii/Pantoea agglomerans* | 4Rz6 | $2.67\times10^{-6}$ | $4.36\times10^{-8}$ | 0.13 | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | 0.02 | 0.81 | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia grimesii* | 9Rz10 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia grimesii/Pantoea agglomerans* | 3Rp10 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia odorifera/Pantoea agglomerans* | 4Rx13 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia plymuthica* | 3Rr8 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia plymuthica* | 4Rc6 | $5.81\times10^{-6}$ | $<1.36\times10^{-8}$ | 0.07 | $2.5\times10^{-8}$ | $7.03\times10^{-6}$ | 0.02 | 0.96 | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia plymuthica* | 4Rx5 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia proteomaculans* | 3Rr9 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia proteomaculans* | 3Rc15 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Weeksella zoohelcum* | 8Rx9 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |

**Table 7.9** continued. Host range of Chi-like phages against rhizosphere isolates library

| Bacteria (determined by FAME) | Strain | Chi-like phages EOPs | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΦLT | ΦRAY0 | ΦRAY1 | ΦRAY19 | ΦRAY26 | ΦRAY27 | ΦJB2 | ΦJB15 | ΦYSD1 |
| *P. luminescens laumondii* | TT01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| *Weeksella zoohelcum* | 5Rr4 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Xenorhabdus luminescens* | 3Rp5 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Xenorhabdus nematophila* | 4Rx4 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | 0.05 | $<1.44\times10^{-8}$ | $2.26\times10^{-6}$ | ND | 0.35 | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Pantoea agglomerans* | 10Bp14 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |
| *Serratia grimesii* | 3Bz10 | $<2.62\times10^{-8}$ | $<1.36\times10^{-8}$ | $<2.18\times10^{-8}$ | $<1.44\times10^{-8}$ | $<4.54\times10^{-8}$ | $<4.76\times10^{-9}$ | $<7.70\times10^{-9}$ | $<2.80\times10^{-8}$ | $<6.23\times10^{-8}$ |

ND - Not determined

## 7.5 Discussion

*Ph. luminescens* has a particularly interesting lifecycle, forming a symbiotic relationship with nematodes but a pathogenic one with insects (Joyce et al., 2006). Briefly, the bacteria live in the gut of the nematode *Heterorhabditis* which provides a safe habitat for them. These nematodes then actively seek out and infect insect larvae. Once infected the nematodes release the *Ph. luminescens* bacteria into the bloodstream where the bacteria release numerous different toxins that kill the insect as well as an antibiotic that prevents growth of other bacteria (Ffrench-Constant et al., 2007; Williams et al., 2005). The nematodes and bacteria are then able to use the insect cadaver as a nutrient source to replicate after which the progeny nematodes carrying the bacteria in their gut leave to seek out new insect hosts to begin the cycle anew. *Ph. luminescens* receives its name as during its replication in the insect larvae it produces a luciferase protein causing the corpse to glow blue (Poinar et al., 1980). This glowing ability is thought to be the cause of the phenomenon called Angel's glow during the battle of Shiloh (Haynes, 2015). During this American Civil war battle, the wounds of certain soldiers began to emit a blue light, these soldiers' wounds would heal faster than those that did not glow which was attributed to divine intervention at the time. It is now thought that this glow was caused by a *Ph. luminescens* infection which would also produce antibiotics preventing infection by pathogenic bacteria and lead to faster wound healing (Haynes, 2015). However, although beneficial to human health in this case, the *Photorhabdus* species is viewed as an emerging human pathogen with increasing reports of human infection (Gerrard et al., 2003). The insecticidal toxin complexes of *Photorhadus* additionally have homologues in *Yersinia pestis* the causative agent of plague which may be a cause of further concern as *Photorhabdus* is regarded as an emerging human pathogen (Rodou et al., 2010).

The remarkable lifecycle of *Ph. luminescens* presents interesting challenges for phages that infect it as the phages too will have to adapt to the different environments that the bacteria cycle through during growth. One way in which phages have adapted is to take shelter within the bacterium itself by resorting to a temperate lifecycle. Indeed this may be the case for many of the phages that infected TT01 in the past as 4% of its genome is associated with prophage DNA (Gaudriault et al., 2004). The work in this chapter displays an alternative means of adaptation. The TT01 phages isolated in this

chapter are capable of infecting a wide range of bacteria originating from very different environments. This strategy to act as a generalist phage with a wide host range allows such phages to have many opportunities to interact with a suitable permissive host and to replicate and maintain viral population. However, the results of the TT01 host range tests also suggest that there are several disadvantages in targeting so many different bacteria. Many of the TT01 phages, after replicating in TT01 then replicate very poorly in some hosts, and for others are unable to infect at all. Some TT01 phages have evolved to no longer be able to infect certain hosts which suggest that they may have gained certain mutations that now allow it to infect other hosts more efficiently. None of the TT01 phages show the same host range patterns implying that these phages may be constantly evolving, depending on the hosts that they may have encountered in the past. Several host range mutants isolated agree with this view as they were shown to replicate just as efficiently in their new host as well as TT01 compared to their wild type counterparts. Although not tested in this chapter due to time constraints it is very likely that the host range profile of these mutants will have changed from their original wild type. One of the remarkable observations was that these TT01 were highly similar in to each other and yet have wide host ranges. This is particularly interesting as the high genomic similarities would make these phages highly susceptible to CRISPR-Cas systems. It remains unknown whether many spacers of different bacteria include the genomes of TT01 phages.

All the TT01 phages isolated were shown to be *Chiviruses* named after phage Chi isolated in 1936 (Sertic and Boulgakov, 1936). The receptor for phage Chi is the flagellum and the phage was shown to infect a wide range of motile bacterial strains including *Salmonella* (Meynell, 1961), *E. coli* (Schade and Adler, 1967) and *Serratia* (Iino and Mitani, 1967). The TT01 phages were only able to infect motile bacteria and so, by targeting a common bacterial component are able to infect many strains of bacteria. It seems likely that phage Chi and other *Chiviruses* will have much larger host ranges than previously thought and could be able to infect many of the strains that the TT01 phage are able to infect. This is supported by the YSD1 phage data. This phage was isolated on a strain of *Salmonella* but was subsequently shown here to be able to infect strains that were also susceptible to TT01 phage infection. Phylogeny trees constructed with the TT01 phages compared to the other *Chiviruses* show that the

TT01 phages do not belong in a specific clade and are just as related to the other *Chiviruses* are to each other (**Figure 7.6**). This result and the fact that the TT01 phages all have different host range profiles suggests they are all replicating on different hosts in the environment. Thus although phages are usually referred to from the hosts they are isolated from *e.g. Salmonella* phages or coliphages, it is perhaps not completely accurate to refer to the TT01 phages as *Photorhabdus* (TT01) phages (Adriaenssens and Brister, 2017). It seems very unlikely that there will be a significant abundance of *Ph. luminescens* growing in the river Cam, hence it is more likely that these "TT01" phages are naturally growing on alternative environmental hosts. The nature of these hosts remains a mystery but it may be far wider than implied by the host range tests. One feature is that none of the TT01 phages are capable of infecting *P. aeruginosa* PA01, although it is motile. Only one strain of *P. aeruginosa* was tested as there were not many strains of *P. aeruginosa* available while this work was in progress but it is likely that if several more strains were tested perhaps one would be susceptible to infection with these phages. Nonetheless, these phages are of particular interest in terms of phage therapy due to their extensive host range in important pathogens. Flagella are important virulence factors in *Salmonella* (Carsiotis et al., 1984) and in *Photorhabdus* play a role in competitive fitness during infection of the insect larvae (Easom and Clarke, 2008).

The mechanism by which phage Chi infects is still poorly understood. One of the models proposed is that Chi infects via a nut and bolt approach (Samuel et al., 1999; Schade et al., 1967). In this model the tail fibre of phage Chi fits into grooves on the surface of the flagella and spins down to the base of the flagellum that is attached to the bacterial surface. The rotating flagellum is thought to provide the force necessary to allow the phage to travel down the flagella in this way. Once at the base, phage Chi is then thought to inject its genomic material. The direction in which the flagella turns is also crucial for successful adsorption as Chi and another *Chivirus* (iEPS5) can only adsorb to bacteria whose flagella are turning in a counter-clockwise direction (Choi et al., 2013). In bacteria with peritrichous flagella, such as *Salmonella* and *E. coli*, the overall direction in which these bacteria move is by a process called run and tumble (Armitage, 2016; Larsen et al., 1974). In this process, when bacteria need to move towards a more favourable location they tend to "run" *i.e.* travel in a straight line more

often when orientated in that direction and “tumble” more until they are correctly orientated. The change from run and tumble is caused by the direction in which the flagella are rotating. When the flagella are turning counter-clockwise the bacteria are motile, and run, but when the flagella are turning clockwise the flagella are not ordered and so the bacteria tumble instead. At least two *Chiviruses* have been shown to be dependent on flagella that are turning counter-clockwise (Choi et al., 2013). Adsorption may occur when certain flagellar residues are exposed when the flagella are rotating in this way or it could follow the nut and bolt theory in which the rotating flagella force the phage to the flagellar base. The nut and bolt theory is likely not the only mechanism through which *Chiviruses* inject their DNA however, as there is evidence the *Chivirus* iESP5 injects its DNA directly into the flagellum and the authors propose that the DNA travels inside the flagellum to reach the bacterial body (Choi et al., 2013). Although the size of DNA may allow it to fit into the narrow central cavity of the flagellum, it is not yet understood how the DNA would bypass the motor components of the flagellum at the base. The phylogenetic trees show that a few of the TT01 *Chiviruses* are more closely related to iESP5 which could suggest they follow a similar mode of infection (**Figure 7.6**). While the adsorption mechanism(s) by which *Chiviruses* bind to flagella and infect their genetic material is unclear, the TT01 phages have been shown to be able to evolve host range mutants when infecting other susceptible hosts. By characterising the mutations that these host range mutant phages evolve it may be possible to identify residues important in adsorption. Furthermore, *Chiviruses* are easy to isolate when using TT01 and so this provides the opportunity to isolate and study even more variants of this genera of phage.

The TT01 phages were originally isolated to see if the $TenpIN_{Pl}$ system located chromosomally in *Ph. luminescens* could abort any TT01 phages. None of the TT01 phages isolated however were aborted, even when the system was introduced in multiple copies via a plasmid. Thus the role of $TenpIN_{Pl}$ within TT01 is still unclear, although with the many different environments TT01 encounters during the replication of its nematode symbiont, it is likely that during certain conditions, the ability to halt bacterial growth or kill a small proportion of the population is advantageous. The $ToxIN_{Bt}$ system has no reported anti-phage activity but functions as a plasmid maintenance system during development of spores (Short et al., 2015) so there may

yet be a function still to be revealed for the TenpIN$_{PI}$ system. The TenpIN$_{PI}$ system was of particular interest as a chromosomal TA system with unknown roles. Many experiments using chromosomal TA systems often highly expresses these TA systems and thus may not define a physiologically relevant functions (Ning et al., 2013; Pedersen et al.). This is also the case with the TenpIN$_{PI}$ system that is shown to have abortive infection properties, but this is only when on a plasmid where it is expressed in much higher amounts than in its natural host TT01 (Blower et al., 2012b). The functions of chromosomal TA systems are still not known however, one possible role in which chromosomal TA systems may be involved is as anti-addiction modules (Bast et al., 2008). `Clearly, further work is required to determine the exact role of the TenpIN$_{PI}$ system in TT01.

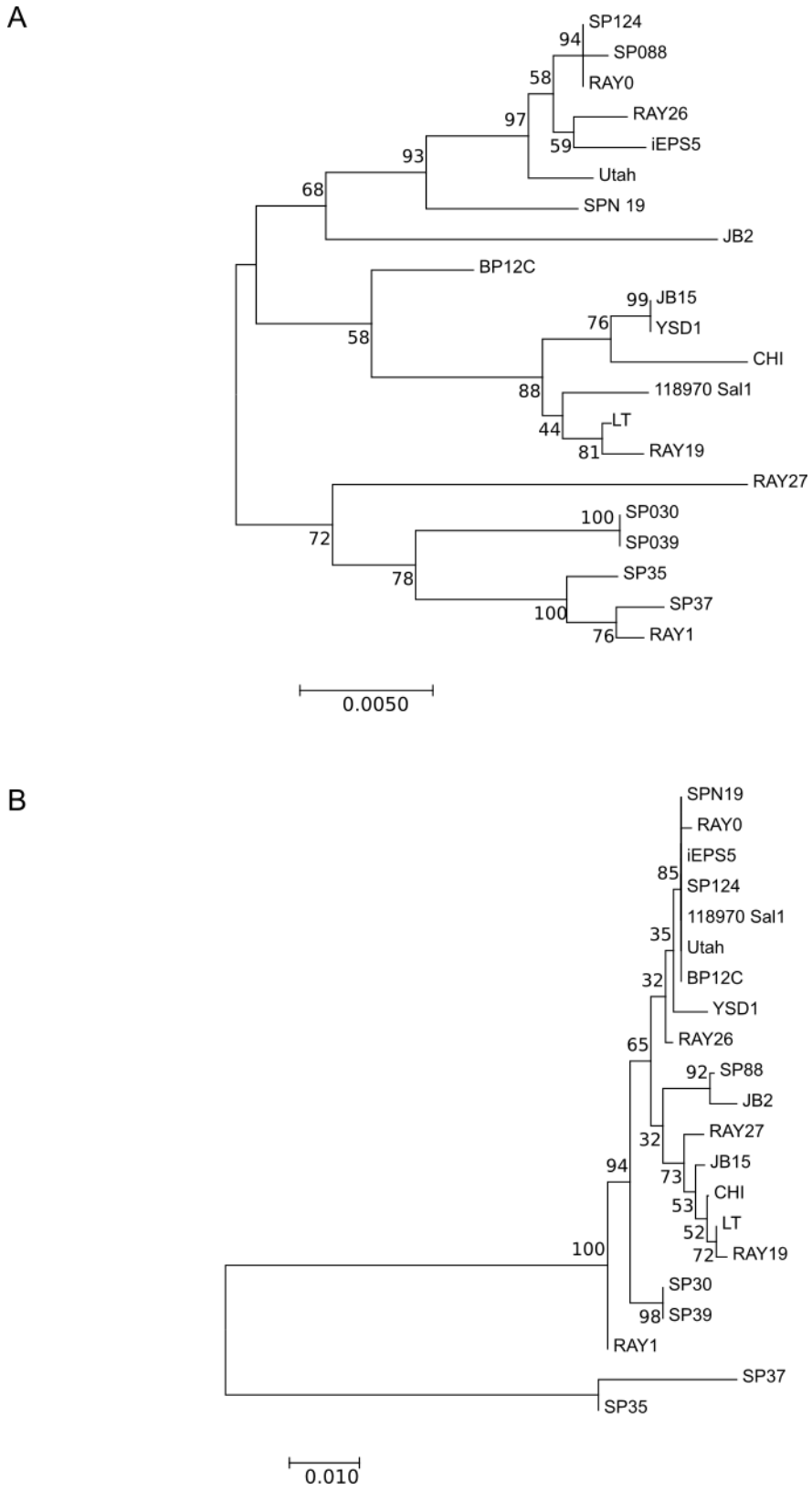



**Figure 7.6** TT01 phages do not fit into a specific clade

Phylogenetic trees were constructed using the neighbour-joining method with a bootstrap test of 1000 replicates. Two different phage proteins were used **A**. Major

capsid protein or **B**. Large terminase. Proteins were aligned using the MUSCLE algorithm.

# Chapter 8 Final discussion

The Type III TA systems have been shown to be potent anti-phage systems aborting a wide range of phages. One of the methodologies used to determine the factors involved in Type III TA activation is to identify the nature of the mutations that phages evolve to escape these systems. Several of these phages, and their mutations, have been identified. However, apart from four phages, all mutations were either found in genes of unknown functions or, in the case of ΦTE, the evolution of an antitoxin mimic. One of the phages with a mutation in a known product is phage 2, which escapes the AbiQ system in *E. coli*, MG1655, through mutations in a putative DNA polymerase. The other three phages are highly similar with three different types of mutations. Only one of the mutations is in a known gene coding for the σ appropriation protein, AsiA. The results described in this dissertation reveal novel phage escape loci and candidate factors related to Type III TA activation. The study begins to build a wider picture of how and why these TA systems are activated during phage infection.

## 8.1 Summary of findings

The first Type III TA system identified was ToxIN$_{Pa}$, which was located on a cryptic plasmid carried by *P. atrosepticum* SCRI1039 and was found to be a potent anti-phage system (Fineran et al., 2009). In this project, to understand ToxIN$_{Pa}$ activation in physiologically natural settings, newly and previously isolated *P. atrosepticum* phages were tested for their susceptibilities to ToxIN$_{Pa}$ using plasmid pTA46. Genomic sequencing of 36 of these *P. atrosepticum* phages revealed six different groups, with four of the groups never seen before among *P. atrosepticum* phages. Highlights of the analysis of the genomic data included several phages showing very high genomic identities to ΦTE, which escapes the ToxIN$_{Pa}$ system by producing an RNA antitoxin mimic (Blower et al., 2012a). These new ΦTE-like phages, however, were unable to escape and were found to either have only a partial antitoxin repeat sequence or none at all, hence they are unable to evolve to produce an antitoxin mimic. These findings also imply that, without the antitoxin mimic sequence, ΦTE itself would be unable to escape the ToxIN$_{Pa}$ system. It is particularly interesting that ΦTE contains a homologue of the DNA polymerase that is mutated in AbiQ escapes of phage 2 yet corresponding mutations were not seen in ΦTE or any of the new phages. None of the newly

sequenced ΦTE-like phages could code for an antitoxin mimic, and this discovery served to highlight the apparent uniqueness of ΦTE and its serendipitous discovery. The only other Abi-sensitive *P. atrosepticum* phage to be characterised in detail is ΦM1 and extensive characterisation of new escapes confirmed that a single protein is responsible for the escape of $ToxIN_{Pa}$, and $TenpIN_{Pl}$. ΦM1 was found to be more sensitive to $ToxIN_{Pa}$ than $TenpIN_{Pl}$. However, the reverse was found in another newly sequenced ΦM1-like phage, ΦRC10. $ToxIN_{Pa}$ and $TenpIN_{Pl}$ escapes of ΦRC10 were also found to have mutations in the homologue of the corresponding protein mutated in ΦM1 escapes. In addition, a new group of *P. atrosepticum* phages, ΦRC6-like, were also found to be aborted by the $ToxIN_{Pa}$ system but were unable to produce escapes. Interestingly, the ΦRC6-like phages carried a putative CRISPR repeat that has not been seen before in any *P. atrosepticum* phages.

The other three groups of phages were not aborted by either of the Type III TA systems used and included the ΦR1-like phages. This group of phages shared high homology to a *P. carotovorum* phage suggesting that ΦR1-like phages may also be able to infect *P. carotovorum*. Indeed, subsequent host range tests of these phages confirmed that certain strains of *P. carotovorum* were indeed susceptible to these phages. This observation led to the idea that at least for the ΦR1-like group, some phages named after the host on which they were isolated, may not necessarily reflect the host on which they replicate in the environment. The final two groups of phages had particularly interesting genomes, as they belonged to a rare class of phages called jumbophages (**Chapter 3**). Many of the proteins from these two phages do not show homology to any previously sequenced gene products and presumably present exciting and completely novel phage biology, yet to be discovered.

$ToxIN_{Pa}$ and $TenpIN_{Pl}$ were previously found to abort many phages infecting *E. coli* DH5α. However, the evolution of escape mutants was not seen before. Coliphage ΦCHAI8 represents the first $ToxIN_{Pa}$-sensitive coliphage that is able to evolve escapes. $ToxIN_{Pa}$ escapes of ΦCHAI8 had deletions in a putative RNA repair locus. Additionally, ΦCHAI8 shared high genomic similarity to phage 2, which is sensitive to the Type III TA system, AbiQ. Escapes of AbiQ have mutations in a putative DNA polymerase. It was a curious result as other than ΦCHAI8, phage 2 was also the only other coliphage that could escape the AbiQ system.

The majority of work performed with ToxIN$_{Pa}$ and AbiQ using coliphages has used either DH5α or MG1655 (Blower et al., 2012b; Samson et al., 2013). However, there is now evidence that, for an unknown reason, certain coliphages that are aborted but can't escape ToxIN$_{Pa}$ or TenpIN$_{Pl}$, are able to do so in the *E. coli* strain EPI300 (**Chapter 6**). This implies that the activation of these Type III TA systems can be also dependent on one or more host factors. One of the phages, T6, that could escape the TenpIN$_{Pl}$ system in EPI300 was characterised further. Several TenpIN$_{Pl}$ escapes of T6 that were sequenced had mutations in the *alc* gene, encoding a transcriptional inhibitor.

The final phage to be characterised with regards to abortive infection was the LamB-dependent *Serratia* phage, ΦOT8s. This phage is aborted by the ToxIN$_{Pa}$ system in both *Serratia* 39006 and *E. coli* DH5α. ToxIN$_{Pa}$ escapes selected from the two hosts have mutations in *motA* and *asiA* genes. This is the second time that a mutant *asiA* gene has been defined as the escape locus in response to ToxIN$_{Pa}$ as prior work in this laboratory showed that three *Serratia* phages, ΦCHI14, ΦX20 and ΦCBH8 also had mutations in *asiA* (Chen et al., 2017).

Finally, the nature of *Photorhabdus* TT01 phages were examined in the last results chapter. All TT01 phages isolated thus far belong to the "new" *Chivirus* genus and have extraordinarily large host ranges. The differences seen in the host range tests suggested that there is constant evolution occurring between these phages depending on their past host histories. None of the TT01 phages were aborted, which could mean that TenpIN$_{Pl}$ has a different natural role in TT01.

In total, forty-five novel phage genomes were sequenced from phages isolated on *P. atrosepticum* (**Chapter 3**), *E. coli* (**Chapter 4**) and *Ph. luminescens* (**Chapter 7**). A taxonomy tree of these phages is shown in **Figure 8.1**. Many of these phages were tested for their susceptibility to the ToxIN$_{Pa}$ or TenpIN$_{Pl}$ systems by transferring the plasmids pTA46 or pFR2, respectively, into the relevant hosts. It is clear that "activation" of both these systems can occur in many different types of phages and in many different hosts.

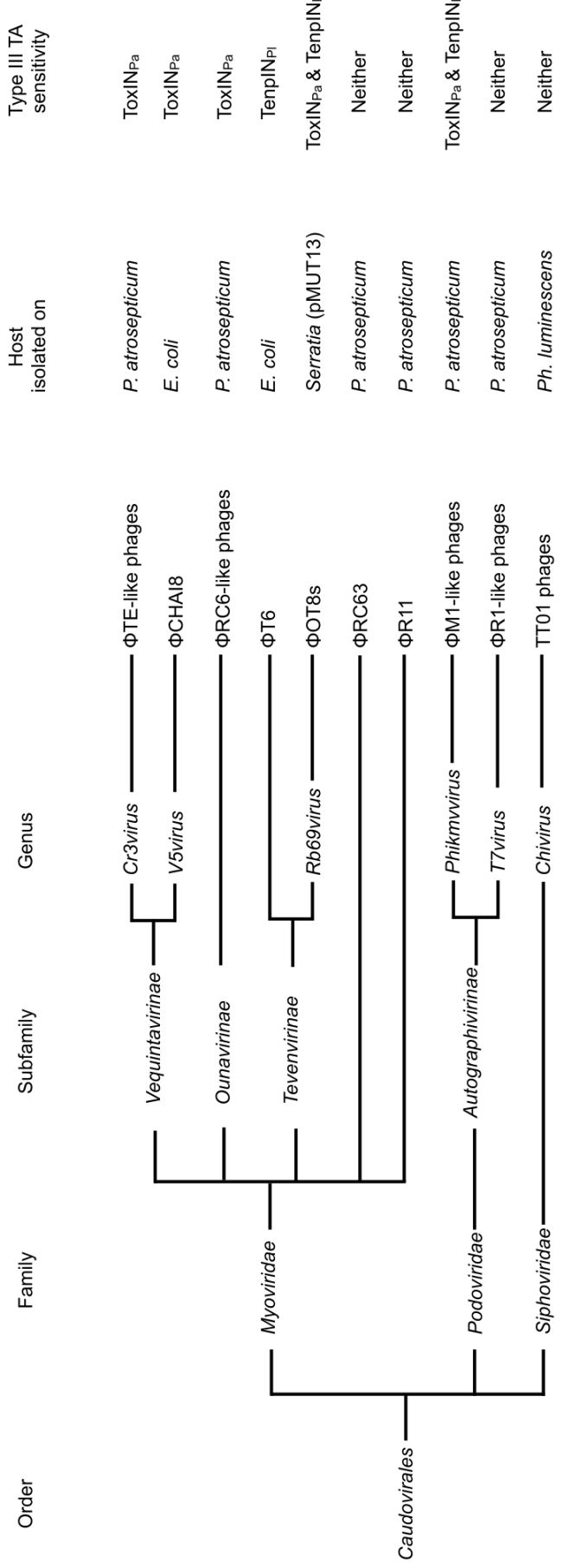



**Figure 8.1** Taxonomy tree of phages used during this study

## 8.2 Proposed model for Type III TA activation

Several new escape loci have been identified in this study. All known escape loci are shown in **Table 8.1**. Many of the identified escape loci carry mutations in single genes of unknown functions. However, it appears that these genes are all within areas associated with nucleotide metabolism. Furthermore, the escape loci that do have known functions are also associated with nucleotide metabolism i.e. processes that affect the transcription of genes in both the phage and their bacterial hosts. Only three different Type III TA systems has been studied in detail but it is clear that all of the three TA systems can be transferred to different bacterial host backgrounds and are still able to abort phages in those hosts (Blower et al., 2012b; Chen et al., 2017; Samson et al., 2013). Perhaps the most extreme transfer would be the AbiQ system originating from a Gram positive i.e. *L. lactis* to a Gram negative i.e. *E. coli* where the phages that infect each host would be widely different.

The escape loci of phage escape mutants of Type III TA systems are associated with nucleotide metabolism (**Table 8.1**). This implies that the factor(s) activating the Type III TA systems is a fundamental process that involves a bacterial-phage interaction linked to the identified escape loci. It is evident that the bacterial host itself plays a key role in activating the Type III TA systems. As different bacterial hosts that carry identical plasmids with the same Type III TA system can behave differently to the same phage. For example, ΦOT8s is aborted by the TenpIN$_{PI}$ system but can only escape in an EPI300 background and not on DH5α or *S* 39006. Moreover, the factor(s) activating Type III TA systems are linked to an essential component of the phage life cycle. Numerous phages aborted by Type III TA systems are unable to escape even when challenged at high titres. However, there are phages that can evolve escapes but this doesn't necessarily mean that this component of the phage cycle is not essential. An explanation for this is that for phages that can escape, this essential can still be performed but perhaps altered in a way that does not lead to Type III TA activation.

**Table 8.1** All known escape loci of Type III TA sensitive phages

| Phage | Host | Type III TA system | Escape loci | Notes |
|---|---|---|---|---|
| ΦTE | *P. atrosepticum* SCRI1043 | ToxIN$_{Pa}$ (pTA46) | ToxI$_{Pa}$ mimic | Ch4 suggests that without ToxI$_{Pa}$ ΦTE can't escape |
| ΦM1/ΦRC10 | *P. atrosepticum* SCRI1043 | ToxIN$_{Pa}$ (pTA46) | M1-23/RC10-26 | Toxic protein of unknown function. Gene located between two nuclease genes. |
| | | TenpIN$_{Pl}$ (pFR2) | M1-23/RC10-26 | Same as above |
| ΦCHAI8 | *E. coli* DH5α | ToxIN$_{Pa}$ (pTA46) | RNA repair | Deletion of several proteins that always includes a putative RNA repair system |
| Phage 2 | *E. coli* MG1655 | AbiQ (pNZ123-AbiQ) | DNA polymerase | Highly similar phage to ΦCHAI8 |
| T6 | *E. coli* EPI300 | TenpIN$_{Pl}$ (pFR2) | *alc* | Several escapes have mutations in a transcriptional inhibitor |
| ΦOT8s | *E. coli* DH5α & *Serratia* ATCC 39006 (pMUT13) | ToxIN$_{Pa}$ (pTA46) | *motA* or *asiA* | Involved in σ factor appropriation |
| | *E. coli* EPI300 | TenpIN$_{Pl}$ (pFR2) | Unknown | Escapes can only be isolated on an EPI300 background |
| ΦCHI14, ΦX20 & ΦCBH8 | *Serratia* ATCC 39006 | ToxIN$_{Pa}$ (pTA46) | Large deletion, *orf84* and *asiA* | Large deletion removes genes encoding various proteins of unknown functions and tRNAs, ORF84 has an unknown function but is an "early" gene and AsiA is involved in sigma appropriation |
| P008 | *L. lactis* IL1403 | AbiQ (pSRQ928) | *orf38* | Unknown function, early gene located in the region involved in DNA replication and nucleotide metabolism |
| bIL170 | *L. lactis* IL1403 | AbiQ (pSRQ925) | *m1* | Unknown function, middle gene located in the region involved in DNA repair and recombination |
| Phage c2 | *L. lactis* LM0230 | AbiQ (pSRQ928) | *e19* | Unknown function, early gene flanked by genes involved in DNA replication |

One process that fits all these criteria (i.e. essential to the phage but mutable to a certain degree, affected by phage/host combination as well as the Type III TA system) is transcription. Every phage is an obligate bacterial parasite as they do not carry all the genes required to replicate independently of their hosts. Consequently, phages have to use the bacterial transcriptional machinery to replicate their genetic material, thus making it an essential process. However, every phage has evolved different means of taking over the host cell machinery and the way that this occurs for the same phage will also be different depending on the host it infects. For example, the process in which the ΦOT8s phage hijacks the bacterial cell machinery will be slightly different when infecting *E. coli* compared to *Serratia*. While the phages have to use the bacterial transcriptional machinery to replicate, certain phage genes involved in this process can be altered, thus changing the way in which the phage hijacks the cell. These phage products may not necessarily be essential, e.g. the M1-23 and RC10-26 proteins of the ΦM1 and ΦRC10 phages can be effectively deleted without any obvious fitness effects to the phage at least under lab conditions (**Chapter 4**).

However, as transcription of phage genes by hijacking the bacterial transcriptional machinery is essential for phage replication, many of the viral genes that are involved in this process will also be essential. This may explain why certain phages are unable to evolve escapes as the mutations that may prevent Type III TA activation, are also be required for taking over the bacterial transcriptional machinery. The mutations of different proteins may affect the rates of transcription differently which explains why selection of a Type III TA system escape mutant on one host will not necessary lead to escape in another host. An example of this is the ΦOT8s phage. A ToxIN$_{Pa}$ escape can be selected on DH5α that is still aborted by the ToxIN$_{Pa}$ system in *Serratia* (**Chapter 7**). The explanation behind this is that as the same phage is exploiting different bacterial transcriptional machinery when infecting the two different hosts, the rate of transcription will be different as well. This may mean that the mutation that allows it to escape ToxIN$_{Pa}$ in DH5α may not have the same impact in *Serratia*. Nonetheless, there are ΦOT8s phage mutations that will have a significant impact in both *E. coli* and *Serratia* such as the *motA* and *asiA* mutations which can then lead to ToxIN$_{Pa}$ escape in both hosts (**Chapter 6**).

The control of transcription is pertinent to the regulation of the Type III TA systems and this is highlighted by the genetic organisation of Type III TA loci. The genetic locus of Type III TA systems is composed of multiple repeats of the antitoxin subunit followed by a Rho-independent terminator that separates it from the toxin sequence. The locus is under the control of a single promoter and with the presence of the terminator this leads to an antitoxin to toxin transcript ratio of ~10:1 for the ToxIN$_{Pa}$ system. Moreover, the ToxI$_{Pa}$ subunit is repeated 5.5 times resulting in 50 antitoxin "units" to a single toxin. The control of antitoxin to toxin stoichiometry is highly important because a decreased level of antitoxin can lead to the activation of the Type III TA system (Fineran et al., 2009). The ratio of antitoxin to toxin is a feature that is also important in other TA systems including Type I and II. The ToxIN$_{Pa}$ system has been shown to form a heterohexameric complex consisting of three antitoxins and three toxins (Blower et al., 2011). However, it has recently been shown that this complex is not the prevalent state of toxin and antitoxin *in vivo*. Instead, a much more dynamic process of neutralisation by the antitoxin and toxin was observed where the ToxN$_{Pa}$ is constantly cleaving newly transcribed antitoxin (Short et al., 2018). Thus it seems entirely plausible that, during phage infection, a generalised inhibition of host translation and/or the promotion of phage replication will lead to a decrease in the level of antitoxin, leading eventually to the release of the toxin. The ΦM1 phage infection results supports this as, during infection, the ToxI$_{Pa}$ levels have been shown to drop faster in the wild type phage compared to an escape mutant (Blower et al., 2017). The ToxN$_{Pa}$ levels do not increase during infection implying that the activation is caused by either the decrease of the antitoxin levels or disruption of the complex. As mentioned above, the disruption of the complex is unlikely since this doesn't seem to be the predominant form *in vivo*. Furthermore, the ΦM1 phage is aborted by both the ToxIN$_{Pa}$ and TenpIN$_{Pl}$ system and escapes are through mutations of the same protein. A single protein is not likely to be able to directly activate these two very different TA systems, thus it is more likely that the protein is in fact affecting a common process determining the activation of the systems. Further evidence that a single phage protein is not likely to directly activate a Type III TA system is evident in phage escapes of the AbiQ system. Both P008 and bIL170 are lactococcal phages that are able to escape the AbiQ system through defects in a single protein. However, while they share homologues of the same escape proteins, the gene that is mutated is different in the two phages (Samson et al., 2013).

ToxIN$_{Pa}$, TenpIN$_{Pl}$ and AbiQ all have different abortive infection profiles although, for ΦM1 and ΦRC10, it seems that a single protein can effect two Type III TA systems. How then does a decrease in transcription efficiency cause certain phages to be aborted by one system but not the other? As mentioned above the levels of antitoxin play a key role in the activation of these systems and in fact there is high variability in the antitoxin sequence (Goeders et al., 2016). During the initial screen for Type III TA systems it became apparent that copy number differed widely between TA systems. Experiments determining the effects of the antitoxin copy number has also shown that there are variabilities in the number of antitoxin repeats necessary to neutralise the toxin. For ToxIN$_{Pa}$, 1.5x ToxI$_{Pa}$ repeats using the native promoter could never be cloned, suggesting that it was insufficient for neutralisation. The smallest repeat that could be cloned was 2.5x, which is in contrast to the results of the AbiQ system where a repeat of only 1.8x was necessary (Bélanger and Moineau, 2015). Additionally, phage mutants of P008 isolated on the AbiQ system with 1.8x repeats did not have any mutations in *orf38*, suggesting that the gene product does not directly activate AbiQ (Bélanger and Moineau, 2015). This agrees with the proposed model that the mutation in these P008 phage escapes must have a more significant impact on the transcriptional rates. Furthermore, a ToxN$_{Pa}$ mutant, ToxN-Y115A, which encodes a toxin that forms a less stable complex with ToxI$_{Pa}$, was still able to abort phages that were previously selected on, and insensitive to, the wild type ToxIN$_{Pa}$ system (Blower et al., 2011; Short et al., 2018). This highlights the importance of the ToxI-ToxN interaction and any disturbance of this (i.e. by decreasing the antitoxin levels), leads to the activation of the Type III TA system.

The ideal antiphage system would target an essential process that is widespread among phages, thereby having as much coverage as possible with limited capacity for resistance to develop. As every phage has to use the transcriptional machinery of their hosts having an antiphage system that is sensitive to transcription, fulfils these criteria (i.e. essential and widespread). Some phages could get around this by evolving a replication cycle that has minimal effect on the bacterial transcriptional machinery, so that these systems aren't activated. Alternatively, other phages may have a fast replication cycle that may be too quick for these systems to detect and respond to. If Type III TA systems activation are reliant on transcription of their antitoxins, it would explain why TA systems from other bacteria could abort phages that infect completely

different bacteria. It is also the likely explanation as to why copy number of these TA systems as well as the number of antitoxin repeats has, such large impacts on phage resistance.

### 8.3 Future directions

The results obtained during this study have shown just how little of the relevant phage biology is understood and there are many avenues which warrant further investigation. In Chapter 3, 36 novel *P. atrosepticum* phages were sequenced, and many shared very little homology to the other published phages. As *P. atrosepticum* phages are already being used to treat potatoes in "phage therapy" (Branston, 2012), it is crucial that much more is understood about these phages. While some basic characterisation has already been performed on these phages, there are many more interesting characterisation that could be followed up with. For example, transduction assays could be conducted. Given that a large proportion of the sequenced phages are similar to ΦTE and ΦM1. It is likely that many of these phages will be transducers, which could be potentially undesirable in phage therapy cocktails. More host range tests could also be performed and the receptor binding proteins for these phages should also be determined. With regards to abortive infection, the nature(s) of the M1-23 and RC10-26 proteins are still unknown. It would be interesting to define the function of these proteins but, given the toxicity of these proteins, it may be easier to first characterise other escape loci identified in this study.

The escape locus of the ΦCHAI8 phage is one of the novel escape loci that may be easier to characterise. Initial follow up experiments should include complementation assays to determine which components are necessary for $ToxIN_{Pa}$ activation. It would also be necessary to determine the natures of the putative RNA repair pathway encoded in the escape loci. Although there is an interest in this RNA repair pathway with regards to $ToxIN_{Pa}$ activation, there are also other reasons to investigate this further. This RNA repair pathway is also involved in another abortive infection system and furthermore, there are homologues of these proteins in eukaryotes which have been shown to be essential proteins (**Chapter 5**). Complementation assays for the escape loci of all other escapes phages such as T6 and ΦOT8s, should also be performed. The escape proteins of these phages, Alc, AsiA and MotA are well characterised in phage T4 but their activities should also be determined. Complementation assays with these proteins could also be performed with T4 homologues. For all phages that were capable of escaping the two Type III TA systems, the amount of antitoxin and toxin during infection of both the wild type and different escapes should be determined. This will help to test whether or not the current

proposed model that depicts decreased antitoxin levels after phage infection leads to abortive infection.

While there is plenty of further work to be done on the products of the escape loci of these phages, the EPI300 experiments have also shown that there are many more escape phages that could be characterised and their escape loci determined (**Chapter 6**). Many of these new phages show abortive infection profiles that are different from any of the currently characterised phages, so are worth of further study.

The final chapter of this dissertation focussed on the phage Chi, a very old phage whose adsorption mechanism has yet to be solved (**Chapter 7**). The phages examined and discussed within the chapter had the potential to create many host range mutants, creating the opportunity to identify residues important in adsorption. This could help to narrow down important components of *Chiviruses*' tail fibres and the definition of the flagellum proteins of the hosts critical for adsorption.

Type III Toxin antitoxin research has come a long way from the original discovery of $ToxIN_{Pa}$ on a cryptic plasmid. Since then, there has been extensive research performed to determine its activation mechanisms. While these are still unknown, the work performed during this research has helped extend our understanding of how Type III TA systems function as abortive infection systems.

## Chapter 9 References

Aakre, C.D., Phung, T.N., Huang, D., and Laub, M.T. (2013). A bacterial toxin inhibits DNA replication elongation through a direct interaction with the β sliding clamp. Mol. Cell *52*, 617–628.

Ackermann, H.-W. (2009). Phage Classification and Characterization. In Bacteriophages, (Humana Press), pp. 127–140.

Adriaenssens, E.M., and Brister, J.R. (2017). How to Name and Classify Your Phage: An Informal Guide. Viruses *9*.

Aiewsakun, P., Adriaenssens, E.M., Lavigne, R., Kropinski, A.M., and Simmonds, P. (2018). Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: steps towards a unified taxonomy. J. Gen. Virol.

Ainsworth, S., Stockdale, S., Bottacini, F., Mahony, J., and van Sinderen, D. (2014). The *Lactococcus lactis* plasmidome: much learnt, yet still lots to discover. FEMS Microbiol. Rev. *38*, 1066–1088.

Amitsur, M., Levitz, R., and Kaufmann, G. (1987). Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA. EMBO J. *6*, 2499–2503.

Amitsur, M., Benjamin, S., Rosner, R., Chapman-Shimshoni, D., Meidler, R., Blanga, S., and Kaufmann, G. (2003). Bacteriophage T4-encoded Stp can be replaced as activator of anticodon nuclease by a normal host cell metabolite. Molecular Microbiology *50*, 129–143.

Arber, W. (1965). Host-Controlled Modification of Bacteriophage. Annual Review of Microbiology *19*, 365–378.

Armitage, J.P. (2016). Classic Spotlight: Seeing Is Believing—Imaging the Active Bacterial Flagellar Filaments. J. Bacteriol. *198*, 2391–2392.

Aucken, H.M., and Pitt, T.L. (1998). Antibiotic resistance and putative virulence factors of *Serratia marcescens* with respect to O and K serotypes. J. Med. Microbiol. *47*, 1105–1113.

Bair, C.L., and Black, L.W. (2007). A Type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. J Mol Biol *366*, 768–778.

Balsalobre, C., Johansson, J., Uhlin, B.E., Rez, A.J., and Oa, F.J.M. (1999). Alterations in Protein Expression Caused by the hha Mutation in *Escherichia coli*: Influence of Growth Medium Osmolarity. J. BACTERIOL. *181*, 7.

Banks, D.J., Lei, B., and Musser, J.M. (2003). Prophage induction and expression of prophage-encoded virulence factors in group A *Streptococcus* serotype M3 strain MGAS315. Infect. Immun. *71*, 7079–7086.

Barford, D., Das, A.K., and Egloff, M.-P. (1998). The structure and mechanism of protein phosphatases: Insights into Catalysis and Regulation. Annual Review of Biophysics and Biomolecular Structure *27*, 133–164.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science *315*, 1709–1712.

Barrios, A.F.G., Zuo, R., Ren, D., and Wood, T.K. (2006). Hha, YbaJ, and OmpA regulate *Escherichia coli* K12 biofilm formation and conjugation plasmids abolish motility. Biotechnol. Bioeng. *93*, 188–200.

Bartual, S.G., Otero, J.M., Garcia-Doval, C., Llamas-Saiz, A.L., Kahn, R., Fox, G.C., and Raaij, M.J. van (2010). Structure of the bacteriophage T4 long tail fiber receptor-binding tip. PNAS *107*, 20287–20292.

Bassler, B.L., Wright, M., and Silverman, M.R. (1994). Multiple signalling systems controlling expression of luminescence in *Vibrio harveyi*: sequence and function of genes encoding a second sensory pathway. Mol. Microbiol. *13*, 273–286.

Bast, M.S.D., Mine, N., and Melderen, L.V. (2008). Chromosomal Toxin-Antitoxin Systems May Act as Antiaddiction Modules. J. Bacteriol. *190*, 4603–4609.

Bebeacua, C., Lorenzo Fajardo, J.C., Blangy, S., Spinelli, S., Bollmann, S., Neve, H., Cambillau, C., and Heller, K.J. (2013). X-ray structure of a superinfection exclusion lipoprotein from phage TP-J34 and identification of the tape measure protein as its target. Mol. Microbiol. *89*, 152–165.

Bélanger, M., and Moineau, S. (2015). Mutational Analysis of the Antitoxin in the Lactococcal Type III Toxin-Antitoxin System AbiQ. Appl. Environ. Microbiol. *81*, 3848–3855.

Bell, K.S., Sebaihia, M., Pritchard, L., Holden, M.T.G., Hyman, L.J., Holeva, M.C., Thomson, N.R., Bentley, S.D., Churcher, L.J.C., Mungall, K., et al. (2004). Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. atroseptica and characterization of virulence factors. PNAS *101*, 11105–11110.

Bennett, H.P.J., and Clarke, D.J. (2005). The pbgPE operon in *Photorhabdus luminescens* Is Required for Pathogenicity and Symbiosis. Journal of Bacteriology *187*, 77–84.

Benzer, S. (1952). Resistance to ultraviolet light as an index to the reproduction of bacteriophage. J Bacteriol *63*, 59–72.

Benzer, S. (1959). On the Topology of the Genetic Fine Structure. PNAS *45*, 1607–1620.

Berg, G., Roskot, N., Steidle, A., Eberl, L., Zock, A., and Smalla, K. (2002). Plant-Dependent Genotypic and Phenotypic Diversity of Antagonistic Rhizobacteria Isolated from Different Verticillium Host Plants. Appl Environ Microbiol *68*, 3328–3338.

Bergh, Ø., BØrsheim, K.Y., Bratbak, G., and Heldal, M. (1989). High abundance of viruses found in aquatic environments. Nature *340*, 467–468.

Bergsland, K.J., Kao, C., Yu, Y.-T.N., Gulati, R., and Snyder, L. (1990). A site in the T4 bacteriophage major head protein gene that can promote the inhibition of all translation in *Escherichia coli*. Journal of Molecular Biology *213*, 477–494.

Bernard, P., Kézdy, K.E., Van Melderen, L., Steyaert, J., Wyns, L., Pato, M.L., Higgins, P.N., and Couturier, M. (1993). The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. J. Mol. Biol. *234*, 534–541.

Bertani, G., and Weigle, J.J. (1953). Host controlled variation in bacterial viruses. J. Bacteriol. *65*, 113–121.

Bertozzi Silva, J., Storms, Z., and Sauvageau, D. (2016). Host receptors for bacteriophage adsorption. FEMS Microbiol Lett *363*.

Bidnenko, E., Ehrlich, S.D., and Chopin, M.C. (1998). *Lactococcus lactis* phage operon coding for an endonuclease homologous to RuvC. Mol. Microbiol. *28*, 823–834.

Bidnenko, E., Chopin, M.-C., Ehrlich, S.D., and Anba, J. (2002). *Lactococcus lactis* AbiD1 abortive infection efficiency is drastically increased by a phage protein. FEMS Microbiol Lett *214*, 283–287.

Bidnenko, E., Chopin, A., Ehrlich, S.D., and Chopin, M.-C. (2009). Activation of mRNA translation by phage protein and low temperature: the case of *Lactococcus lactis* abortive infection system AbiD1. BMC Molecular Biology *10*, 4.

Bigger, J. (1944). Treatment of Staphylococcal infections with penicillin by intermittent sterilisation. The Lancet *244*, 497–500.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., et al. (1997). The complete genome sequence of *Escherichia coli* K-12. Science *277*, 1453–1462.

Blower, T.R. (2009). Functional and structural studies of the toxIN abortive infection and toxin-antitoxin locus from *Erwinia carotovora* subspecies atroseptica. University of Cambridge.

Blower, T.R., Fineran, P.C., Johnson, M.J., Toth, I.K., Humphreys, D.P., and Salmond, G.P.C. (2009). Mutagenesis and Functional Characterization of the RNA and Protein Components of the toxIN Abortive Infection and Toxin-Antitoxin Locus of *Erwinia*. J. Bacteriol. *191*, 6029–6039.

Blower, T.R., Pei, X.Y., Short, F.L., Fineran, P.C., Humphreys, D.P., Luisi, B.F., and Salmond, G.P.C. (2011). A processed noncoding RNA regulates an altruistic bacterial antiviral system. Nature Structural & Molecular Biology *18*, 185–190.

Blower, T.R., Short, F.L., Rao, F., Mizuguchi, K., Pei, X.Y., Fineran, P.C., Luisi, B.F., and Salmond, G.P.C. (2012b). Identification and classification of bacterial Type III toxin–antitoxin systems encoded in chromosomal and plasmid genomes. Nucleic Acids Res *40*, 6158–6173.

Blower, T.R., Evans, T.J., Przybilski, R., Fineran, P.C., and Salmond, G.P.C. (2012a). Viral Evasion of a Bacterial Suicide System by RNA–Based Molecular Mimicry Enables Infectious Altruism. PLOS Genetics *8*, e1003023.

Blower, T.R., Chai, R., Przybilski, R., Chindhy, S., Fang, X., Kidman, S.E., Tan, H., Luisi, B.F., Fineran, P.C., and Salmond, G.P.C. (2017). Evolution of *Pectobacterium* Bacteriophage ΦM1 To Escape Two Bifunctional Type III Toxin-Antitoxin and Abortive Infection Systems through Mutations in a Single Viral Gene. Appl. Environ. Microbiol. *83*, e03229-16.

Boemare, N.E., Akhurst, R.J., and Mourant, R.G. (1993). DNA Relatedness between *Xenorhabdus* spp. (Enterobacteriaceae), Symbiotic Bacteria of Entomopathogenic Nematodes, and a Proposal To Transfer *Xenorhabdus luminescens* to a New Genus, *Photorhabdus* gen. nov. International Journal of Systematic Bacteriology *43*, 249–255.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology (Reading, Engl.) *151*, 2551–2561.

Bonocora, R.P., Caignan, G., Woodrell, C., Werner, M.H., and Hinton, D.M. A basic/hydrophobic cleft of the T4 activator MotA interacts with the C-terminus of *E. coli* σ70 to activate middle gene transcription. Molecular Microbiology *69*, 331–343.

Botstein, D., and Herskowitz, I. (1974). Properties of hybrids between *Salmonella* phage P22 and coliphage lambda. Nature *251*, 584–589.

Branston (2012). A natural solution to takle potential soft rot | Branston Limited.

Brantl, S., and Jahn, N. (2015). sRNAs in bacterial type I and type III toxin-antitoxin systems. FEMS Microbiol Rev *39*, 413–427.

Breitbart, M. (2011). Marine Viruses: Truth or Dare. Annu. Rev. Mar. Sci. *4*, 425–448.

Brody, H., and Hill, C.W. (1988). Attachment site of the genetic element e14. J. Bacteriol. *170*, 2040–2044.

Brüssow, H., and Hendrix, R.W. (2002). Phage genomics: small is beautiful. Cell *108*, 13–16.

Brzozowska, I., and Zielenkiewicz, U. (2013). Regulation of toxin–antitoxin systems by proteolysis. Plasmid *70*, 33–41.

Burstein, D., Sun, C.L., Brown, C.T., Sharon, I., Anantharaman, K., Probst, A.J., Thomas, B.C., and Banfield, J.F. (2016). Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. Nature Communications *7*, 10613.

Buttimer, C., McAuliffe, O., Ross, R.P., Hill, C., O'Mahony, J., and Coffey, A. (2017a). Bacteriophages and Bacterial Plant Diseases. Front Microbiol *8*.

Buttimer, C., Hendrix, H., Oliveira, H., Casey, A., Neve, H., McAuliffe, O., Ross, R.P., Hill, C., Noben, J.-P., O'Mahony, J., et al. (2017b). Things Are Getting Hairy: Enterobacteria Bacteriophage vB_PcaM_CBB. Front Microbiol *8*.

Buttimer, C., Born, Y., Lucid, A., Loessner, M.J., Fieseler, L., and Coffey, A. (2018). *Erwinia amylovora* phage vB_EamM_Y3 represents another lineage of hairy Myoviridae. Research in Microbiology.

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the Bacterial RNA Polymerase Promoter Specificity σ Subunit. Molecular Cell *9*, 527–539.

Carsiotis, M., Weinstein, D.L., Karch, H., Holder, I.A., and O'Brien, A.D. (1984). Flagella of *Salmonella typhimurium* are a virulence factor in infected C57BL/6J mice. Infect Immun *46*, 814–818.

Casjens, S.R., and Hendrix, R.W. (2015). Bacteriophage lambda: Early pioneer and still relevant. Virology *479–480*, 310–330.

Cataudella, I., Trusina, A., Sneppen, K., Gerdes, K., and Mitarai, N. (2012). Conditional cooperativity in toxin-antitoxin regulation prevents random toxin activation and promotes fast translational recovery. Nucleic Acids Res. *40*, 6424–6434.

Chan, W.T., Espinosa, M., and Yeo, C.C. (2016). Keeping the Wolves at Bay: Antitoxins of Prokaryotic Type II Toxin-Antitoxin Systems. Front. Mol. Biosci. *3*.

Chen, B., Akusobi, C., Fang, X., and Salmond, G.P.C. (2017). Environmental T4-Family Bacteriophages Evolve to Escape Abortive Infection via Multiple Routes in a Bacterial Host Employing “Altruistic Suicide” through Type III Toxin-Antitoxin Systems. Front Microbiol *8*.

Cheng, H.-Y., Soo, V.W.C., Islam, S., McAnulty, M.J., Benedik, M.J., and Wood, T.K. (2014). Toxin GhoT of the GhoT/GhoS toxin/antitoxin system damages the cell membrane to reduce adenosine triphosphate and to reduce growth under stress. Environ. Microbiol. *16*, 1741–1754.

Cherny, I., Rockah, L., and Gazit, E. (2005). The YoeB toxin is a folded protein that forms a physical complex with the unfolded YefM antitoxin. Implications for a structural-based differential stability of toxin-antitoxin systems. J. Biol. Chem. *280*, 30063–30072.

Choi, J., Kotay, S.M., and Goel, R. (2010). Various physico-chemical stress factors cause prophage induction in *Nitrosospira multiformis* 25196--an ammonia oxidizing bacteria. Water Res. *44*, 4550–4558.

Choi, Y., Shin, H., Lee, J.-H., and Ryu, S. (2013). Identification and Characterization of a Novel Flagellum-Dependent *Salmonella*-Infecting Bacteriophage, iEPS5. Appl. Environ. Microbiol. *79*, 4829–4837.

Chopin, M.-C., Chopin, A., and Bidnenko, E. (2005). Phage abortive infection in lactococci: variations on a theme. Current Opinion in Microbiology *8*, 473–479.

Chowdhury, N., Kwan, B.W., and Wood, T.K. (2016). Persistence Increases in the Absence of the Alarmone Guanosine Tetraphosphate by Reducing Cell Growth. Sci Rep *6*, 20519.

Chowdhury, R., Biswas, S.K., and Das, J. (1989). Abortive replication of choleraphage phi 149 in *Vibrio cholerae* biotype el tor. J. Virol. *63*, 392–397.

Christensen-Dalsgaard, M., Jørgensen, M.G., and Gerdes, K. (2010). Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. Mol. Microbiol. *75*, 333–348.

Christie, G.E., and Dokland, T. (2012). Pirates of the Caudovirales. Virology *434*, 210–221.

Cicero, M.P., Alexander, K.A., and Kreuzer, K.N. (1998). The MotA Transcriptional Activator of Bacteriophage T4 Binds to Its Specific DNA Site as a Monomer. Biochemistry *37*, 4977–4984.

Clark, J.R., and March, J.B. (2006). Bacteriophages and biotechnology: vaccines, gene therapy and antibacterials. Trends in Biotechnology *24*, 212–218.

Clokie, M.R., Millard, A.D., Letarov, A.V., and Heaphy, S. (2011). Phages in nature. Bacteriophage *1*, 31–45.

Colland, F., Orsini, G., Brody, E.N., Buc, H., and Kolb, A. (1998a). The bacteriophage T4 AsiA protein: a molecular switch for sigma 70-dependent promoters. Mol. Microbiol. *27*, 819–829.

Colland, F., Orsini, G., Brody, E.N., Buc, H., and Kolb, A. (1998b). The bacteriophage T4 AsiA protein: a molecular switch for sigma 70-dependent promoters. Mol. Microbiol. *27*, 819–829.

Comeau, A.M., Bertrand, C., Letarov, A., Tétart, F., and Krisch, H.M. (2007). Modular architecture of the T4 phage superfamily: A conserved core genome and a plastic periphery. Virology *362*, 384–396.

Connelly, J.C., and Leach, D.R.F. (2002). Tethering on the brink: the evolutionarily conserved Mre11–Rad50 complex. Trends in Biochemical Sciences *27*, 410–418.

Copeland, N.A., and Kleanthous, C. (2005). The Role of an Activating Peptide in Protease-mediated Suicide of Escherichia coli K12. J. Biol. Chem. *280*, 112–117.

Crick, F.H., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961). General nature of the genetic code for proteins. Nature *192*, 1227–1232.

Cruz, J.W., Rothenbacher, F.P., Maehigashi, T., Lane, W.S., Dunham, C.M., and Woychik, N.A. (2014). Doc toxin is a kinase that inactivates elongation factor Tu. J. Biol. Chem. *289*, 7788–7798.

Czyz, A., Los, M., Wrobel, B., and Wegrzyn, G. (2001). Inhibition of spontaneous induction of lambdoid prophages in *Escherichia coli* cultures: simple procedures with possible biotechnological applications. BMC Biotechnol. *1*, 1.

Depew, R.E., and Cozzarelli, N.R. (1974). Genetics and Physiology of Bacteriophage T4 3′-Phosphatase: Evidence for Involvement of the Enzyme in T4 DNA Metabolism. J. Virol. *13*, 888–897.

Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. J. Bacteriol. *190*, 1390–1400.

D'Herelle, F. (2007). On an invisible microbe antagonistic toward dysenteric bacilli: brief note by Mr. F. D'Herelle, presented by Mr. Roux. Research in Microbiology *158*, 553–554.

Dörr, T., Vulić, M., and Lewis, K. (2010). Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. PLoS Biol. *8*, e1000317.

Drivdahl, R.H., and Kutter, E.M. (1990). Inhibition of transcription of cytosine-containing DNA in vitro by the alc gene product of bacteriophage T4. J Bacteriol *172*, 2716–2727.

Duckworth, D.H. (1976). "Who discovered bacteriophage?". Bacteriol Rev *40*, 793–802.

Dy, R.L., Richter, C., Salmond, G.P.C., and Fineran, P.C. (2014a). Remarkable Mechanisms in Microbes to Resist Phage Infections. Annual Review of Virology *1*, 307–331.

Dy, R.L., Przybilski, R., Semeijn, K., Salmond, G.P.C., and Fineran, P.C. (2014b). A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism. Nucleic Acids Res. *42*, 4590–4605.

Easom, C.A., and Clarke, D.J. (2008). Motility is required for the competitive fitness of entomopathogenic *Photorhabdus luminescens* during insect infection. BMC Microbiol *8*, 168.

Ellis, E.L., and Delbrück, M. (1939). The growth of bacteriophage. J Gen Physiol *22*, 365–384.

Emond, E., Dion, E., Walker, S.A., Vedamuthu, E.R., Kondo, J.K., and Moineau, S. (1998). AbiQ, an abortive infection mechanism from *Lactococcus lactis*. Appl. Environ. Microbiol. *64*, 4748–4756.

Englert, M., and Beier, H. (2005). Plant tRNA ligases are multifunctional enzymes that have diverged in sequence and substrate specificity from RNA ligases of other phylogenetic origins. Nucleic Acids Res. *33*, 388–399.

Ffrench-Constant, R.H., Dowling, A., and Waterfield, N.R. (2007). Insecticidal toxins from *Photorhabdus* bacteria and their potential use in agriculture. Toxicon *49*, 436–451.

Fineran, P.C., Blower, T.R., Foulds, I.J., Humphreys, D.P., Lilley, K.S., and Salmond, G.P.C. (2009). The phage abortive infection system, ToxIN, functions as a protein–RNA toxin–antitoxin pair. PNAS *106*, 894–899.

Finnin, M.S., Cicero, M.P., Davies, C., Porter, S.J., White, S.W., and Kreuzer, K.N. (1997). The activation domain of the MotA transcription factor from bacteriophage T4. The EMBO Journal *16*, 1992–2003.

Fischer-Le Saux, M., Viallard, V., Brunel, B., Normand, P., and Boemare, N.E. (1999). Polyphasic classification of the genus *Photorhabdus* and proposal of new taxa: *P. luminescens* subsp. luminescens subsp. nov., *P. luminescens* subsp. akhurstii subsp. nov., *P. luminescens* subsp. laumondii subsp. nov., *P. temperata* sp. nov., *P. temperata* subsp. temperata subsp. nov. and *P. asymbiotica* sp. nov. Int. J. Syst. Bacteriol. *49 Pt 4*, 1645–1656.

Flyg, C., Kenne, K., and Boman, H.G. (1980). Insect pathogenic properties of *Serratia marcescens*: phage-resistant mutants with a decreased resistance to Cecropia immunity and a decreased virulence to Drosophila. J. Gen. Microbiol. *120*, 173–181.

Fortier, L.-C., and Sekulovic, O. (2013). Importance of prophages to evolution and virulence of bacterial pathogens. Virulence *4*, 354–365.

Fozo, E.M., Makarova, K.S., Shabalina, S.A., Yutin, N., Koonin, E.V., and Storz, G. (2010). Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. Nucleic Acids Res. *38*, 3743–3759.

Fruciano, D.E., and Bourne, S. (2007). Phage as an antimicrobial agent: d'Herelle's heretical theories and their role in the decline of phage prophylaxis in the West. Can J Infect Dis Med Microbiol *18*, 19–26.

Furnham, N., Holliday, G.L., de Beer, T.A.P., Jacobsen, J.O.B., Pearson, W.R., and Thornton, J.M. (2014). The Catalytic Site Atlas 2.0: cataloging catalytic sites and residues identified in enzymes. Nucleic Acids Res *42*, D485–D489.

García-Contreras, R., Zhang, X.-S., Kim, Y., and Wood, T.K. (2008). Protein translation and cell death: the role of rare tRNAs in biofilm formation and in activating dormant phage killer genes. PLoS ONE *3*, e2394.

Garneau, J.E., and Moineau, S. (2011). Bacteriophages of lactic acid bacteria and their impact on milk fermentations. Microbial Cell Factories *10*, S20.

Garvey, P., Fitzgerald, G.F., and Hill, C. (1995). Cloning and DNA sequence analysis of two abortive infection phage resistance determinants from the lactococcal plasmid pNP40. Applied and Environmental Microbiology *61*, 4321–4328.

Gaudriault, S., Thaler, J.-O., Duchaud, E., Kunst, F., Boemare, N., and Givaudan, A. (2004). Identification of a P2-related prophage remnant locus of *Photorhabdus luminescens* encoding an R-type phage tail-like particle. FEMS Microbiol. Lett. *233*, 223–231.

Gebhart, D., Williams, S.R., Bishop-Lilly, K.A., Govoni, G.R., Willner, K.M., Butani, A., Sozhamannan, S., Martin, D., Fortier, L.-C., and Scholl, D. (2012). Novel high-molecular-weight, R-type bacteriocins of *Clostridium difficile*. J. Bacteriol. *194*, 6240–6247.

George, D.G., Yeh, L.-S.L., and Barker, W.C. (1983). Unexpected relationships between bacteriophage lambda hypothetical proteins and bacteriophage T4 tail-fiber proteins. Biochemical and Biophysical Research Communications *115*, 1061–1068.

Georgiades, K., and Raoult, D. (2011). Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. PLoS ONE *6*, e17962.

Gerdes, K., Rasmussen, P.B., and Molin, S. (1986a). Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. PNAS *83*, 3116–3120.

Gerdes, K., Bech, F.W., Jørgensen, S.T., Løbner-Olesen, A., Rasmussen, P.B., Atlung, T., Boe, L., Karlstrom, O., Molin, S., and von Meyenburg, K. (1986b). Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the *E. coli* relB operon. EMBO J. *5*, 2023–2029.

Gerrard, J.G., McNevin, S., Alfredson, D., Forgan-Smith, R., and Fraser, N. (2003). *Photorhabdus* Species: Bioluminescent Bacteria as Human Pathogens? Emerg Infect Dis *9*, 251–254.

Ghazal, S., Oshone, R., Simpson, S., Morris, K., Abebe-Akele, F., Thomas, W.K., Khalil, K.M., and Tisa, L.S. (2016). Draft Genome Sequence of *Photorhabdus luminescens* subsp. laumondii HP88, an Entomopathogenic Bacterium Isolated from Nematodes. Genome Announc *4*.

Goeders, N., Chai, R., Chen, B., Day, A., and Salmond, G.P.C. (2016). Structure, Evolution, and Functions of Bacterial Type III Toxin-Antitoxin Systems. Toxins *8*, 282.

Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A.C., and Kuriyan, J. (1995). Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. Nature *376*, 745–753.

Goormaghtigh, F., Fraikin, N., Putrinš, M., Hallaert, T., Hauryliuk, V., Garcia-Pino, A., Sjödin, A., Kasvandik, S., Udekwu, K., Tenson, T., et al. (2018). Reassessing the Role of Type II Toxin-Antitoxin Systems in Formation of *Escherichia coli* Type II Persister Cells. MBio *9*, e00640-18.

Gottesman, M.E., and Weisberg, R.A. (2004). Little Lambda, Who Made Thee? Microbiol. Mol. Biol. Rev. *68*, 796–813.

Griffith, J.P., Kim, J.L., Kim, E.E., Sintchak, M.D., Thomson, J.A., Fitzgibbon, M.J., Fleming, M.A., Caron, P.R., Hsiao, K., and Navia, M.A. (1995). X-Ray Structure of Calcineurin Inhibited by the Immunophilin-Immunosuppressant FKBP12-FK506 Complex. 16.

Gross, S.R. (1954). Abortive Infection of a Strain of *Escherichia Coli* by Coliphage T2. J. Bacteriol. *68*, 36–42.

Guglielmini, J., and Van Melderen, L. (2011). Bacterial toxin-antitoxin systems: Translation inhibitors everywhere. Mob Genet Elements *1*, 283–290.

Guo, Y., Quiroga, C., Chen, Q., McAnulty, M.J., Benedik, M.J., Wood, T.K., and Wang, X. (2014). RalR (a DNase) and RalA (a small RNA) form a type I toxin-antitoxin system in *Escherichia coli*. Nucleic Acids Res. *42*, 6448–6462.

Hallez, R., Geeraerts, D., Sterckx, Y., Mine, N., Loris, R., and Van Melderen, L. (2010). New toxins homologous to ParE belonging to three-component toxin-antitoxin systems in *Escherichia coli* O157:H7. Mol. Microbiol. *76*, 719–732.

Hammad, A.M.M. (1998). Evaluation of alginate-encapsulated *Azotobacter chroococcum* as a phage-resistant and an effective inoculum. Journal of Basic Microbiology *38*, 9–16.

Hampton, H.G., Jackson, S.A., Fagerlund, R.D., Vogel, A.I.M., Dy, R.L., Blower, T.R., and Fineran, P.C. (2018). AbiEi Binds Cooperatively to the Type IV abiE Toxin-Antitoxin Operator Via a Positively-Charged Surface and Causes DNA Bending and Negative Autoregulation. J. Mol. Biol. *430*, 1141–1156.

Han, Q., Zhou, C., Wu, S., Liu, Y., Triplett, L., Miao, J., Tokuhisa, J., Deblais, L., Robinson, H., Leach, J.E., et al. (2015). Crystal Structure of Xanthomonas AvrRxo1-ORF1, a Type III Effector with a Polynucleotide Kinase Domain, and Its Interactor AvrRxo1-ORF2. Structure *23*, 1900–1909.

Hanson, P.I., and Whiteheart, S.W. (2005). AAA+ proteins: have engine, will work. Nature Reviews Molecular Cell Biology *6*, 519–529.

Harms, A., Fino, C., Sørensen, M.A., Semsey, S., and Gerdes, K. (2017). Prophages and Growth Dynamics Confound Experimental Results with Antibiotic-Tolerant Persister Cells. MBio *8*, e01964-17.

Harms, A., Brodersen, D.E., Mitarai, N., and Gerdes, K. (2018). Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. Molecular Cell *70*, 768–784.

Harrison, E., and Brockhurst, M.A. (2017). Ecological and Evolutionary Benefits of Temperate Phage: What Does or Doesn't Kill You Makes You Stronger. BioEssays *39*, 1700112.

Hashemolhosseini, S., Stierhof, Y.D., Hindennach, I., and Henning, U. (1996). Characterization of the helper proteins for the assembly of tail fibers of coliphages T4 and lambda. J. Bacteriol. *178*, 6258–6265.

Haynes, A. (2015). Glowing wounds and angelic bacteria | Blog.

Hazan, R., and Engelberg-Kulka, H. (2004). *Escherichia coli mazEF* mediated cell death as a defense mechanism that inhibits the spread of phage P1. Mol Genet Genomics *272*, 227–234.

Hendrix, R.W., and Duda, R.L. (1992). Bacteriophage lambda PaPa: not the mother of all lambda phages. Science *258*, 1145–1148.

Herdendorf, T.J., Albrecht, D.W., Benkovic, S.J., and Nelson, S.W. (2011). Biochemical Characterization of Bacteriophage T4 Mre11-Rad50 Complex. J. Biol. Chem. *286*, 2382–2392.

D'Herelle, F. (1918). Technique de la recherche du microbe filtrant bacteriophage (Bacteriophagum intestinale). 1160–1162.

Hill, C.W., Gray, J.A., and Brody, H. (1989). Use of the isocitrate dehydrogenase structural gene for attachment of e14 in *Escherichia coli* K-12. J. Bacteriol. *171*, 4083–4084.

Hirata, H., Kashihara, M., Horiike, T., Suzuki, T., Dohra, H., Netsu, O., and Tsuyumu, S. (2016). Genome Sequence of *Pectobacterium carotovorum* Phage PPWS1,

Isolated from Japanese Horseradish [*Eutrema japonicum* (Miq.) Koidz] Showing Soft-Rot Symptoms. Genome Announc *4*.

Hu, Y., Benedik, M.J., and Wood, T.K. (2012). Antitoxin DinJ influences the general stress response through transcript stabilizer CspE. Environ. Microbiol. *14*, 669–679.

Hua, J. (2016). Capsid Structure and DNA Packing in Jumbo Bacteriophages. PhD Thesis. University of Pittsburgh.

Hynes, A.P., Rousseau, G.M., Agudelo, D., Goulet, A., Amigues, B., Loehr, J., Romero, D.A., Fremaux, C., Horvath, P., Doyon, Y., et al. (2018). Widespread anti-CRISPR proteins in virulent bacteriophages inhibit a range of Cas9 proteins. Nature Communications *9*, 2919.

Iino, T., and Mitani, M. (1967). Infection of *Serratia marcescens* by bacteriophage chi. J Virol *1*, 445–447.

Izard, T. (2001). Structural basis for chloramphenicol tolerance in *Streptomyces venezuelae* by chloramphenicol phosphotransferase activity. Protein Sci *10*, 1508–1513.

Jabbar, M.A., and Snyder, L. (1984). Genetic and physiological studies of an *Escherichia coli* locus that restricts polynucleotide kinase- and RNA ligase-deficient mutants of bacteriophage T4. J. Virol. *51*, 522–529.

Jansen, R., Embden, J.D.A. van, Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. Mol. Microbiol. *43*, 1565–1575.

Jayaraman, R. (2008). Seymour Benzer and T4 rII: Running the map into the ground. Resonance *13*, 898–908.

Joyce, S.A., Watson, R.J., and Clarke, D.J. (2006). The regulation of pathogenicity and mutualism in *Photorhabdus*. Current Opinion in Microbiology *9*, 127–132.

Jurenaite, M., Markuckas, A., and Suziedeliene, E. (2013). Identification and characterization of type II toxin-antitoxin systems in the opportunistic pathogen *Acinetobacter baumannii*. J. Bacteriol. *195*, 3165–3172.

Kalischuk, M., Hachey, J., and Kawchuk, L. (2015). Complete Genome Sequence of Phytopathogenic *Pectobacterium atrosepticum* Bacteriophage Peat1. Genome Announc *3*.

Kawano, M., Aravind, L., and Storz, G. (2007). An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. Mol. Microbiol. *64*, 738–754.

Keen, E.C. (2015). A century of phage research: Bacteriophages and the shaping of modern biology. Bioessays *37*, 6–9.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols *10*, 845–858.

Kim, J.-S., and Wood, T.K. (2016). Persistent Persister Misperceptions. Front Microbiol *7*.

Kim, M.S., Kim, Y.D., Hong, S.S., Park, K., Ko, K.S., and Myung, H. (2015). Phage-Encoded Colanic Acid-Degrading Enzyme Permits Lytic Phage Infection of a Capsule-Forming Resistant Mutant *Escherichia coli* Strain. Applied and Environmental Microbiology *81*, 900–909.

Koga, M., Otsuka, Y., Lemire, S., and Yonesaki, T. (2011). *Escherichia coli* rnlA and rnlB Compose a Novel Toxin–Antitoxin System. Genetics *187*, 123–130.

Krogh, B.O., Llorente, B., Lam, A., and Symington, L.S. (2005). Mutations in Mre11 Phosphoesterase Motif I That Impair *Saccharomyces cerevisiae* Mre11-Rad50-Xrs2 Complex Stability in Addition to Nuclease Activity. Genetics *171*, 1561–1570.

Kropinski, A.M., Waddell, T., Meng, J., Franklin, K., Ackermann, H.-W., Ahmed, R., Mazzocco, A., Yates, J., Lingohr, E.J., and Johnson, R.P. (2013). The host-range, genomics and proteomics of *Escherichia coli* O157:H7 bacteriophage rV5. Virology Journal *10*, 76.

Kuhn, J., Kropinski, A., Anany, H., Tolstoy, I., Kutter, E., and Adriaenssens, E. (2017). To create a new bacteriophage family, *Ackermannviridae*, containing two (2) new subfamilies including four (4) genera.

Kutter, E., Beug, A., Sluss, R., Jensen, L., and Bradley, D. (1975). The production of undegraded cytosine-containing DNA by bacteriophage T4 in the absence of dCTPase and endonucleases II and IV, and its effects on T4-directed protein synthesis. Journal of Molecular Biology *99*, 591–607.

Kutter, E.M., Bradley, D., Schenck, R., Guttman, B.S., and Laiken, R. (1981). Bacteriophage T4 alc gene product: general inhibitor of transcription from cytosine-containing DNA. J Virol *40*, 822–829.

Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. Nature Reviews Microbiology *8*, 317–327.

Landsmann, J., Kröger, M., and Hobom, G. (1982). The rex region of bacteriophage lambda: two genes under three-way control. Gene *20*, 11–24.

Larrañaga, O., Brown-Jaque, M., Quirós, P., Gómez-Gómez, C., Blanch, A.R., Rodríguez-Rubio, L., and Muniesa, M. (2018). Phage particles harboring antibiotic resistance genes in fresh-cut vegetables and agricultural soil. Environ Int *115*, 133–141.

Larsen, S.H., Reader, R.W., Kort, E.N., Tso, W.-W., and Adler, J. (1974). Change in direction of flagellar rotation is the basis of the chemotactic response in Escherichia coli. Nature *249*, 74–77.

Lavigne, R., Burkal'tseva, M.V., Robben, J., Sykilinda, N.N., Kurochkina, L.P., Grymonprez, B., Jonckx, B., Krylov, V.N., Mesyanzhinov, V.V., and Volckaert, G. (2003). The genome of bacteriophage φKMV, a T7-like virus infecting *Pseudomonas aeruginosa*. Virology *312*, 49–59.

Lederberg, E.M., and Lederberg, J. (1953). Genetic Studies of Lysogenicity in *Escherichia Coli*. Genetics *38*, 51.

Lee, C., Kim, J., Shin, S.G., and Hwang, S. (2006). Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. Journal of Biotechnology *123*, 273–280.

Lee, J.-H., Shin, H., Kim, H., and Ryu, S. (2011). Complete Genome Sequence of *Salmonella* Bacteriophage SPN3US. J. Virol. *85*, 13470–13471.

Lehman, I.R., and Pratt, E.A. (1960). On the Structure of the Glucosylated Hydroxymethylcytosine Nucleotides of Coliphages T2, T4, and T6. *235*, 7.

Lehnherr, H., Maguin, E., Jafri, S., and Yarmolinsky, M.B. (1993). Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. J. Mol. Biol. *233*, 414–428.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Drèze, P., and Van Melderen, L. (2011). Diversity of bacterial type II toxin–antitoxin systems: a comprehensive search and functional analysis of novel families. Nucleic Acids Res *39*, 5513–5525.

Lewis, K. (2010). Persister cells. Annu. Rev. Microbiol. *64*, 357–372.

Li, N., Sickmier, E.A., Zhang, R., Joachimiak, A., and White, S.W. The MotA transcription factor from bacteriophage T4 contains a novel DNA-binding domain: the 'double wing' motif. Molecular Microbiology *43*, 1079–1088.

Lima-Mendez, G., Toussaint, A., and Leplae, R. (2007). Analysis of the phage sequence space: The benefit of structured information. Virology *365*, 241–249.

Liu, Q., and C Richardson, C. (1993). Liu, Q. & Richardson, C. C. Gene 5.5 protein of bacteriophage T7 inhibits the nucleoid protein H-NS of *Escherichia coli*. Proc. Natl Acad. Sci. USA 90, 1761-1765. Proceedings of the National Academy of Sciences of the United States of America *90*, 1761–1765.

Łobocka, M.B., Rose, D.J., Plunkett, G., Rusin, M., Samojedny, A., Lehnherr, H., Yarmolinsky, M.B., and Blattner, F.R. (2004). Genome of Bacteriophage P1. J. Bacteriol. *186*, 7032–7068.

Loenen, W.A.M., Dryden, D.T.F., Raleigh, E.A., Wilson, G.G., and Murray, N.E. (2014). Highlights of the DNA cutters: a short history of the restriction enzymes. Nucleic Acids Res *42*, 3–19.

Lohse, D., Denu, J., and Dixon, J. (1995). Insights derived from the structures of the Ser/Thr phosphatases calcineurin and protein phosphatase 1. Structure *3*, 987–990.

Lu, M.J., and Henning, U. (1989). The immunity (imm) gene of *Escherichia coli* bacteriophage T4. J Virol *63*, 3472–3478.

Lu, M.-J., and Henning, U. (1994). Superinfection exclusion by T-even-type coliphages. Trends in Microbiology *2*, 137–139.

Luria, S.E., and Human, M.L. (1952). A nonhereditary, host-induced variation of bacterial viruses. J. Bacteriol. *64*, 557–569.

Luria, S.E., Delbrück, M., and Anderson, T.F. (1943). Electron Microscope Studies of Bacterial Viruses. J. Bacteriol. *46*, 57–77.

Magnuson, R., Lehnherr, H., Mukhopadhyay, G., and Yarmolinsky, M.B. (1996). Autoregulation of the Plasmid Addiction Operon of Bacteriophage P1. J. Biol. Chem. *271*, 18705–18710.

Mai-Prochnow, A., Hui, J.G.K., Kjelleberg, S., Rakonjac, J., McDougald, D., and Rice, S.A. (2015). 'Big things in small packages: the genetics of filamentous phage and effects on fitness of their host.' FEMS Microbiol Rev *39*, 465–487.

Maisonneuve, E., Castro-Camargo, M., and Gerdes, K. (2013). RETRACTED: (p)ppGpp Controls Bacterial Persistence by Stochastic Induction of Toxin-Antitoxin Activity. Cell *154*, 1140–1150.

Maisonneuve, E., Castro-Camargo, M., and Gerdes, K. (2018). Retraction for Maisonneuve et al., Bacterial persistence by RNA endonucleases. PNAS 201803278.

Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2009). Comprehensive comparative-genomic analysis of Type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. Biology Direct *4*, 19.

Marimon, O., Teixeira, J.M.C., Cordeiro, T.N., Soo, V.W.C., Wood, T.L., Mayzel, M., Amata, I., García, J., Morera, A., Gay, M., et al. (2016). An oxygen-sensitive toxin-antitoxin system. Nat Commun *7*, 13634.

Masuda, H., Tan, Q., Awano, N., Wu, K.-P., and Inouye, M. (2012). YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*. Mol. Microbiol. *84*, 979–989.

Matos, R.C., Lapaque, N., Rigottier-Gois, L., Debarbieux, L., Meylheuc, T., Gonzalez-Zorn, B., Repoila, F., Lopes, M. de F., and Serror, P. (2013). *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits. PLoS Genet. *9*, e1003539.

Mayr-Harting, A. (1955). The Acquisition of Penicillin Resistance by *Staphylococcus aureus*, Strain Oxford. Journal of General Microbiology *13*, 9–21.

McClean, K.H., Chhabra, S.R., Camara, M., Daykin, M., Swift, S., Bycroft, B.W., and Williams, P. (2018). Quorum sensing and *Chronobacterium violaceum*:exploitation of violacein production and inhibition for the detection of N-acylhomoserine lactones. 9.

McGowan, S.J., Barnard, A.M.L., Bosgelmez, G., Sebaihia, M., Simpson, N.J.L., Thomson, N.R., Todd, D.E., Welch, M., Whitehead, N.A., and Salmond, G.P.C. (2005). Carbapenem antibiotic biosynthesis in *Erwinia carotovora* is regulated by physiological and genetic factors modulating the quorum sensing-dependent control pathway. Molecular Microbiology *55*, 526–545.

Meyer, J.R., Dobias, D.T., Weitz, J.S., Barrick, J.E., Quick, R.T., and Lenski, R.E. (2012). Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda. Science *335*, 428–432.

Meynell, E.W. (1961). A phage, Φx, which attacks motile bacteria. 41.

Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., and Rüger, W. (2003). Bacteriophage T4 Genome. Microbiol. Mol. Biol. Rev. *67*, 86–156.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. *60*, 174–182.

Monson, R.E., Jones, K., and Salmond, G.P.C. (2018). Draft Genome Sequence of *Pectobacterium carotovorum* subsp. carotovorum ATCC 39048, a Carbapenem-Producing Phytopathogen. Microbiol Res Announc *7*, e00825-18.

Montag, D., Schwarz, H., and Henning, U. (1989). A component of the side tail fiber of *Escherichia coli* bacteriophage lambda can functionally replace the receptor-recognizing part of a long tail fiber protein of the unrelated bacteriophage T4. J. Bacteriol. *171*, 4378–4384.

Montag, D., Hashemolhosseini, S., and Henning, U. (1990). Receptor-recognizing proteins of T-even type bacteriophages: The receptor-recognizing area of proteins 37 of phages T4 TuIa and TuIb. Journal of Molecular Biology *216*, 327–334.

Morona, R., Klose, M., and Henning, U. (1984). Escherichia coli K-12 outer membrane protein (OmpA) as a bacteriophage receptor: analysis of mutant genes expressing altered proteins. J Bacteriol *159*, 570–578.

Naser, I.B., Hoque, M.M., Nahid, M.A., Tareq, T.M., Rocky, M.K., and Faruque, S.M. (2017). Analysis of the CRISPR-Cas system in bacteriophages active on epidemic strains of *Vibrio cholerae* in Bangladesh. Scientific Reports *7*, 14880.

Nelson, D. (2004). Phage Taxonomy: We Agree To Disagree. J. Bacteriol. *186*, 7029–7031.

Nieto, J.M., Madrid, C., Miquelay, E., Parra, J.L., Rodriguez, S., and Juarez, A. (2002). Evidence for Direct Protein-Protein Interaction between Members of the Enterobacterial Hha/YmoA and H-NS Families of Proteins. Journal of Bacteriology *184*, 629–635.

Ning, D., Jiang, Y., Liu, Z., and Xu, Q. (2013). Characterization of a Chromosomal Type II Toxin–Antitoxin System mazEaFa in the *Cyanobacterium Anabaena* sp. PCC 7120. PLOS ONE *8*, e56035.

Noble, R., and Fuhrman, J. (1998). Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. Aquatic Microbial Ecology *14*, 113–118.

Otsuka, Y., and Yonesaki, T. (2012). Dmd of bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnlA toxins. Mol. Microbiol. *83*, 669–681.

Ouhammouch, M., Adelman, K., Harvey, S.R., Orsini, G., and Brody, E.N. (1995). Bacteriophage T4 MotA and AsiA proteins suffice to direct *Escherichia coli* RNA polymerase to initiate transcription at T4 middle promoters. PNAS *92*, 1451–1455.

Overgaard, M., Borch, J., and Gerdes, K. (2009). RelB and RelE of *Escherichia coli* form a tight complex that represses transcription via the ribbon-helix-helix motif in RelB. J. Mol. Biol. *394*, 183–196.

Page, R., and Peti, W. (2016). Toxin-antitoxin systems in bacterial growth arrest and persistence. Nature Chemical Biology *12*, 208–214.

Paget, M.S., and Helmann, J.D. (2003). The σ70 family of sigma factors. Genome Biol *4*, 203.

Pal, D., Vuthoori, M., Pande, S., Wheeler, D., and Hinton, D.M. (2003). Analysis of Regions within the Bacteriophage T4 AsiA Protein Involved in its Binding to the σ70 Subunit of *E.coli* RNA Polymerase and its Role as a Transcriptional Inhibitor and Co-activator. Journal of Molecular Biology *325*, 827–841.

Pande, S., Makela, A., Dove, S.L., Nickels, B.E., Hochschild, A., and Hinton, D.M. (2002). The Bacteriophage T4 Transcription Activator MotA Interacts with the Far-C-Terminal Region of the σ70 Subunit of *Escherichia coli* RNA Polymerase. J. Bacteriol. *184*, 3957–3964.

Pandey, D.P., and Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res *33*, 966–976.

Parma, D.H., Snyder, M., Sobolevski, S., Nawroz, M., Brody, E., and Gold, L. (1992). The Rex system of bacteriophage lambda: tolerance and altruistic cell death. Genes Dev. *6*, 497–510.

Paull, T.T., and Deshpande, R.A. (2014). The Mre11/Rad50/Nbs1 Complex: recent insights into catalytic activities and ATP-driven conformational changes. Exp Cell Res *329*, 139–147.

Pearson, R.E., and Snyder, L. (1980). Shutoff of lambda gene expression by bacteriophage T4: role of the T4 alc gene. J Virol *35*, 194–202.

Pecota, D.C., and Wood, T.K. (1996). Exclusion of T4 phage by the hok/sok killer locus from plasmid R1. J. Bacteriol. *178*, 2044–2050.

Pedersen, K., Christensen, S.K., and Gerdes, K. Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. Molecular Microbiology *45*, 501–510.

Penner, M., Morad, I., Snyder, L., and Kaufmann, G. (1995). Phage T4-coded Stp: Double-Edged Effector of Coupled DNA and tRNA-Restriction Systems. Journal of Molecular Biology *249*, 857–868.

Poinar, G.O., Thomas, G., Haygood, M., and Nealson, K.H. (1980). Growth and luminescence of the symbolic bacteria associated with the terrestrial nematode, *Heterorphabditis bacteriophora*. Soil Biology and Biochemistry *12*, 5–10.

Poinar, G.O., Hess, R.T., Lanier, W., Kinney, S., and White, J.H. (1989). Preliminary observations of a bacteriophage infecting *Xenorhabdus luminescens* (Enterobacteriaceae). Experientia *45*, 191–192.

Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology (Reading, Engl.) *151*, 653–663.

Pritchard, L., Humphris, S., Baeyen, S., Maes, M., Van Vaerenbergh, J., Elphinstone, J., Saddler, G., and Toth, I. (2013a). Draft Genome Sequences of Four *Dickeya dianthicola* and Four *Dickeya solani* Strains. Genome Announc *1*.

Pritchard, L., Humphris, S., Saddler, G.S., Elphinstone, J.G., Pirhonen, M., and Toth, I.K. (2013b). Draft Genome Sequences of 17 Isolates of the Plant Pathogenic Bacterium *Dickeya*. Genome Announc *1*.

Raleigh, E.A., and Wilson, G. (1986). *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. Proc Natl Acad Sci U S A *83*, 9070–9074.

Ramage, H.R., Connolly, L.E., and Cox, J.S. (2009). Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. PLoS Genet. *5*, e1000767.

Rand, K.N., and Gait, M.J. (1984). Sequence and cloning of bacteriophage T4 gene 63 encoding RNA ligase and tail fibre attachment activities. EMBO J *3*, 397–402.

Rao, F. (2014). Structural and functional characterisation of novel Type III systems. University of Cambridge.

Rao, F., Short, F.L., Voss, J.E., Blower, T.R., Orme, A.L., Whittaker, T.E., Luisi, B.F., and Salmond, G.P.C. (2015). Co-evolution of quaternary organization and novel RNA tertiary interactions revealed in the crystal structure of a bacterial protein-RNA toxin-antitoxin system. Nucleic Acids Res. *43*, 9529–9540.

Rath, D., Amlinger, L., Rath, A., and Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. Biochimie *117*, 119–128.

Regué, M., Fabregat, C., and Viñas, M. (1991). A generalized transducing bacteriophage for *Serratia marcescens*. Res. Microbiol. *142*, 23–27.

Reiter, T.A., Reiter, N.J., and Rusnak, F. (2002). $Mn^{2+}$ Is a Native Metal Ion Activator for Bacteriophage λ Protein Phosphatase †. Biochemistry *41*, 15404–15409.

Rifat, D., Wright, N.T., Varney, K.M., Weber, D.J., and Black, L.W. (2008). Restriction endonuclease inhibitor IPI* of bacteriophage T4. J Mol Biol *375*, 720–734.

Ripp, S., and Miller, R.V. (1998). Dynamics of the pseudolysogenic response in slowly growing cells of *Pseudomonas aeruginosa*. Microbiology (Reading, Engl.) *144 ( Pt 8)*, 2225–2232.

Roberts, R.J. (2005). How restriction enzymes became the workhorses of molecular biology. PNAS *102*, 5905–5908.

Robinson, A., Brzoska, A.J., Turner, K.M., Withers, R., Harry, E.J., Lewis, P.J., and Dixon, N.E. (2010). Essential Biological Processes of an Emerging Pathogen: DNA Replication, Transcription, and Cell Division in Acinetobacter spp. Microbiol. Mol. Biol. Rev. *74*, 273–297.

Rodou, A., Ankrah, D.O., and Stathopoulos, C. (2010). Toxins and Secretion Systems of *Photorhabdus luminescens*. Toxins (Basel) *2*, 1250–1264.

S. Makarova, K., H. Haft, D., Barrangou, R., J. J. Brouns, S., Charpentier, E., Horvath, P., Moineau, S., J. M. Mojica, F., I. Wolf, Y., Yakunin, A.F., et al. (2011). Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol *9*, 467–477.

Samson, J.E., Bélanger, M., and Moineau, S. (2013). Effect of the Abortive Infection Mechanism and Type III Toxin/Antitoxin System AbiQ on the Lytic Cycle of *Lactococcus lactis* Phages. J. Bacteriol. *195*, 3947–3956.

Samuel, A.D., Pitta, T.P., Ryu, W.S., Danese, P.N., Leung, E.C., and Berg, H.C. (1999). Flagellar determinants of bacterial sensitivity to chi-phage. Proc. Natl. Acad. Sci. U.S.A. *96*, 9863–9866.

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology *32*, 347–355.

Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F., and Petersen, G.B. (1982). Nucleotide sequence of bacteriophage λ DNA. Journal of Molecular Biology *162*, 729–773.

Santos, S.B., Kropinski, A.M., Ceyssens, P.-J., Ackermann, H.-W., Villegas, A., Lavigne, R., Krylov, V.N., Carvalho, C.M., Ferreira, E.C., and Azeredo, J. (2011). Genomic and Proteomic Characterization of the Broad-Host-Range *Salmonella* Phage PVP-SE1: Creation of a New Phage Genus. J. Virol. *85*, 11265–11273.

Sapriel, G., Wandersman, C., and Delepelaire, P. (2003). The SecB Chaperone Is Bifunctional in *Serratia marcescens*: SecB Is Involved in the Sec Pathway and Required for HasA Secretion by the ABC Transporter. Journal of Bacteriology *185*, 80–88.

Schade, S.Z., and Adler, J. (1967). Purification and Chemistry of Bacteriophage x. J. Virol. *1*, 8.

Schade, S.Z., Adler, J., and Ris, H. (1967). How Bacteriophage χ Attacks Motile Bacteria. J. Virol. *1*, 599–609.

Schicklmaier, P., Moser, E., Wieland, T., Rabsch, W., and Schmieger, H. (1998). A comparative study on the frequency of prophages among natural isolates of *Salmonella* and *Escherichia coli* with emphasis on generalized transducers. Antonie Van Leeuwenhoek *73*, 49–54.

Schuster, H., Beyersmann, D., Mikolajczyk, M., and Schlicht, M. (1973). Prophage Induction by High Temperature in Thermosensitive dna Mutants Lysogenic for Bacteriophage Lambda. J Virol *11*, 879–885.

Schwer, B., Sawaya, R., Ho, C.K., and Shuman, S. (2004). Portability and fidelity of RNA-repair systems. PNAS *101*, 2788–2793.

Schwer, B., Aronova, A., Ramirez, A., Braun, P., and Shuman, S. (2008). Mammalian 2′,3′ cyclic nucleotide phosphodiesterase (CNP) can function as a tRNA splicing enzyme in vivo. RNA (New York, N.Y.) *14*, 204–210.

Seed, K.D., Lazinski, D.W., Calderwood, S.B., and Camilli, A. (2013). A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. Nature *494*, 489–491.

Sertic, V., and Boulgakov, N. (1936). Bactériophages spécifiques pour des variétés bactériennes flagellées. Compt Rend Soc Biol 887–888.

Setlow, J.K., Boling, M.E., Allison, D.P., and Beattie, K.L. (1973). Relationship Between Prophage Induction and Transformation in *Haemophilus influenzae*. J. Bacteriol. *115*, 153–161.

Severinov, K., Kashlev, M., Severinova, E., Bass, I., McWilliams, K., Kutter, E., Nikiforov, V., Snyder, L., and Goldfarb, A. (1994). A non-essential domain of *Escherichia coli* RNA polymerase required for the action of the termination factor Alc. J. Biol. Chem. *269*, 14254–14259.

Shan, Y., Lazinski, D., Rowe, S., Camilli, A., and Lewis, K. (2015). Genetic basis of persister tolerance to aminoglycosides in *Escherichia coli*. MBio *6*.

Sharma, R.C., and Smith, K.C. (1987). Role of DNA polymerase I in postreplication repair: a reexamination with *Escherichia coli* delta polA. J. Bacteriol. *169*, 4559–4564.

Shinedling, S., Parma, D., and Gold, L. (1987). Wild-type bacteriophage T4 is restricted by the lambda rex genes. J Virol *61*, 3790–3794.

Shmakov, S.A., Sitnik, V., Makarova, K.S., Wolf, Y.I., Severinov, K.V., and Koonin, E.V. (2017). The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. MBio *8*, e01397-17.

Short, F.L., Pei, X.Y., Blower, T.R., Ong, S.-L., Fineran, P.C., Luisi, B.F., and Salmond, G.P.C. (2013). Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot. Proc. Natl. Acad. Sci. U.S.A. *110*, E241-249.

Short, F.L., Monson, R.E., and Salmond, G.P.C. (2015). A Type III protein-RNA toxin-antitoxin system from *Bacillus thuringiensis* promotes plasmid retention during spore development. RNA Biol *12*, 933–937.

Short, F.L., Akusobi, C., Broadhurst, W.R., and Salmond, G.P.C. (2018). The bacterial Type III toxin-antitoxin system, ToxIN, is a dynamic protein-RNA complex with stability-dependent antiviral abortive infection activity. Sci Rep *8*.

Shub, D.A. (1994). Bacterial Viruses: Bacterial altruism? Current Biology *4*, 555–556.

Sidhu, S.S. (2001). Engineering M13 for phage display. Biomolecular Engineering *18*, 57–63.

Sirotkin, K., Cooley, W., Runnels, J., and Snyder, L.R. (1978). A role in true-late gene expression for the T4 bacteriophage 5′ polynucleotide kinase 3′ phosphatase. Journal of Molecular Biology *123*, 221–233.

Slavcev, R.A., and Hayes, S. (2002). Rex-Centric Mutualism. J Bacteriol *184*, 857–858.

Smith, H.S., Pizer, L.I., Pylkas, L., and Lederberg, S. Abortive Infection of *Shigella dysenteriae* P2 by T2 Bacteriophage. J. VIROL. 7.

Smith, J.L., Goldberg, J.M., and Grossman, A.D. (2014). Complete Genome Sequences of *Bacillus subtilis* subsp. subtilis Laboratory Strains JH642 (AG174) and AG1839. Genome Announc *2*.

Snyder, L., and Jorissen, L. (1986). Molecular Proof that Bacteriophage T4 alc and unf Genes Are the Same Genet. J. Bacteriol. *168*, 6.

Snyder, L., and McWilliams, K. (1989). The rex genes of bacteriophage lambda can inhibit cell function without phage superinfection. Gene *81*, 17–24.

Snyder, L., Gold, L., and Kutter, E. (1976). A gene of bacteriophage T4 whose product prevents true late transcription on cytosine-containing T4 DNA. Proc Natl Acad Sci U S A *73*, 3098–3102.

Song, S., and Wood, T.K. (2018). Post-segregational Killing and Phage Inhibition Are Not Mediated by Cell Death Through Toxin/Antitoxin Systems. Front. Microbiol. *9*.

Soo, V.W.C., and Wood, T.K. (2013). Antitoxin MqsA represses curli formation through the master biofilm regulator CsgD. Sci Rep *3*, 3186.

Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR — a widespread system that provides acquired resistance against phages in bacteria and archaea. Nature Reviews Microbiology *6*, 181–186.

Staals, R.H.J., Jackson, S.A., Biswas, A., Brouns, S.J.J., Brown, C.M., and Fineran, P.C. (2016). Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native CRISPR–Cas system. Nature Communications *7*, 12853.

Stern, A., and Sorek, R. (2011). The phage-host arms-race: Shaping the evolution of microbes. Bioessays *33*, 43–51.

Steward, G.F., Montiel, J.L., and Azam, F. (2000). Genome size distributions indicate variability and similarities among marine viral assemblages from diverse environments. Limnology and Oceanography *45*, 1697–1706.

Steward, G.F., Culley, A.I., Mueller, J.A., Wood-Charlson, E.M., Belcaid, M., and Poisson, G. (2013). Are we missing half of the viruses in the ocean? The ISME Journal *7*, 672.

Stone, R. (2002). Stalin's Forgotten Cure. Science *298*, 728–731.

Strathern, A., and Herskowitz, I. (1975). Defective prophage in *Escherichia coli* K12 strains. Virology *67*, 136–143.

Summers, W.C. (2001). Bacteriophage Therapy. Annual Review of Microbiology *55*, 437–451.

Suttle, C.A., Chan, A.M., and Cottrell, M.T. (1990). Infection of phytoplankton by viruses and reduction of primary productivity. Nature *347*, 467–469.

Tan, D., Svenningsen, S.L., and Middelboe, M. (2015). Quorum Sensing Determines the Choice of Antiphage Defense Strategy in *Vibrio anguillarum*. MBio *6*.

Tan, Q., Awano, N., and Inouye, M. (2011). YeeV is an *Escherichia coli* toxin that inhibits cell division by targeting the cytoskeleton proteins, FtsZ and MreB. Mol. Microbiol. *79*, 109–118.

Tao, P., Wu, X., and Rao, V. (2018). Unexpected evolutionary benefit to phages imparted by bacterial CRISPR-Cas9. Science Advances *4*, eaar4134.

Tétart, F., Repoila, F., Monod, C., and Krisch, H.M. (1996). Bacteriophage T4 Host Range is Expanded by Duplications of a Small Domain of the Tail Fiber Adhesin. Journal of Molecular Biology *258*, 726–731.

Thisted, T., Sørensen, N.S., Wagner, E.G., and Gerdes, K. (1994). Mechanism of post-segregational killing: Sok antisense RNA interacts with Hok mRNA via its 5'-end single-stranded leader and competes with the 3'-end of Hok mRNA for binding to the mok translational initiation region. EMBO J. *13*, 1960–1968.

Tomaru, Y., and Nagasaki, K. (2007). Flow cytometric detection and enumeration of DNA and RNA viruses infecting marine eukaryotic microalgae. J Oceanogr *63*, 215–221.

Toth, I.K., Mulholland, V., Cooper, V., Bentley, S., Shih, Y.-L., Perombelon, M.C.M., and Salmond, G.P.C. (1997). Generalized transduction in the potato blackleg pathogen *Erwinia carotovora* subsp. atroseptica by bacteriophage M1. Microbiology *143*, 2433–2438.

Tran, L.S., Szabó, L., Ponyi, T., Orosz, L., Sík, T., and Holczinger, A. (1999). Phage abortive infection of *Bacillus licheniformis* ATCC 9800; identification of the abiBL11 gene and localisation and sequencing of its promoter region. Appl. Microbiol. Biotechnol. *52*, 845–852.

Tripathi, A., Dewan, P.C., Siddique, S.A., and Varadarajan, R. (2014). MazF-induced Growth Inhibition and Persister Generation in Escherichia coli. J. Biol. Chem. *289*, 4191–4205.

Triplett, L.R., Shidore, T., Long, J., Miao, J., Wu, S., Han, Q., Zhou, C., Ishihara, H., Li, J., Zhao, B., et al. (2016). AvrRxo1 Is a Bifunctional Type III Secreted Effector and Toxin-Antitoxin System Component with Homologs in Diverse Environmental Contexts. PLOS ONE *11*, e0158856.

Trojet, S.N., Caumont-Sarcos, A., Perrody, E., Comeau, A.M., and Krisch, H.M. (2011). The gp38 Adhesins of the T4 Superfamily: A Complex Modular Determinant of the Phage's Host Specificity. Genome Biol Evol *3*, 674–686.

Turgeon, N., Laflamme, C., Ho, J., and Duchaine, C. (2008). Evaluation of the plasmid copy number in *B. cereus spores*, during germination, bacterial growth and sporulation using real-time PCR. Plasmid *60*, 118–124.

Twort, F.W. (1915). An investigation on the nature of ultra-microscopic viruses. The Lancet *186*, 1241–1243.

V. Shaburova a, O., Hertveldt b, K., Krylov, S., A. Pleteneva a, E., V. Bourkaltseva a, M., Lavigne, R., Volckaert b, G., N. Krylov, V., and Papa, D.M. (2006). Comparison of New Giant Bacteriophages OBP and Lu11 of Soil Pseudomonads with Bacteriophages of the f KZ-Supergroup of *Pseudomonas aeruginosa*. Russian Journal of Genetics *42*, 877–885.

Van Melderen, L., and Saavedra De Bast, M. (2009). Bacterial Toxin–Antitoxin Systems: More Than Selfish Entities? PLoS Genet *5*.

Vovis, G.F., and Lacks, S. (1977). Complementary action of restriction enzymes endo R · DpnI and endo R · DpnII on bacteriophage f1 DNA. Journal of Molecular Biology *115*, 525–538.

Wang, L.K., and Shuman, S. (2002). Mutational analysis defines the 5′-kinase and 3′-phosphatase active sites of T4 polynucleotide kinase. Nucleic Acids Res *30*, 1073–1080.

Wang, L.K., and Shuman, S. (2005). Structure–function analysis of yeast tRNA ligase. RNA *11*, 966–975.

Wang, J., Michel, V., Hofnung, M., and Charbit, A. (1998). Cloning of the J gene of bacteriophage lambda, expression and solubilization of the J protein: first in vitro studies on the interactions between J and LamB, its cell surface receptor. Research in Microbiology *149*, 611–624.

Wang, L.K., Lima, C.D., and Shuman, S. (2002). Structure and mechanism of T4 polynucleotide kinase: an RNA repair enzyme. EMBO J *21*, 3873–3880.

Wang, X., Kim, Y., Hong, S.H., Ma, Q., Brown, B.L., Pu, M., Tarone, A.M., Benedik, M.J., Peti, W., Page, R., et al. (2011). Antitoxin MqsA helps mediate the bacterial general stress response. Nature Chemical Biology *7*, 359–366.

Wang, X., Lord, D.M., Cheng, H.-Y., Osbourne, D.O., Hong, S.H., Sanchez-Torres, V., Quiroga, C., Zheng, K., Herrmann, T., Peti, W., et al. (2012). A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. Nat. Chem. Biol. *8*, 855–861.

Wang, X., Lord, D.M., Hong, S.H., Peti, W., Benedik, M.J., Page, R., and Wood, T.K. (2013). Type II toxin/antitoxin MqsR/MqsA controls type V toxin/antitoxin GhoT/GhoS. Environ. Microbiol. *15*, 1734–1744.

Wei, Y., Chesne, M.T., Terns, R.M., and Terns, M.P. (2015). Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in *Streptococcus thermophilus*. Nucleic Acids Res *43*, 1749–1758.

Weinbauer, M.G. (2004). Ecology of prokaryotic viruses. FEMS Microbiol Rev *28*, 127–181.

Wen, W., Liu, B., Xue, L., Zhu, Z., Niu, L., and Sun, B. (2018). Autoregulation and Virulence Control by the Toxin-Antitoxin System SavRS in *Staphylococcus aureus*. Infect. Immun. *86*.

Williams, J.S., Thomas, M., and Clarke, D.J. (2005). The gene stlA encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in *Photorhabdus luminescens* TT01. Microbiology *151*, 2543–2550.

Williams, R.S., Moncalian, G., Williams, J.S., Yamada, Y., Limbo, O., Shin, D.S., Groocock, L.M., Cahill, D., Hitomi, C., Guenther, G., et al. (2008). Mre11 Dimers Coordinate DNA End Bridging and Nuclease Processing in Double-Strand-Break Repair. Cell *135*, 97–109.

Wittebole, X., De Roock, S., and Opal, S.M. (2014). A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. Virulence *5*, 226–235.

Wood, W.B., and Henninger, M. (1969). Attachment of tail fibers in bacteriophage T4 assembly: Some properties of the reaction in vitro and its genetic control. Journal of Molecular Biology *39*, 603–618.

Yamaguchi, Y., Park, J.-H., and Inouye, M. (2009). MqsR, a crucial regulator for quorum sensing and biofilm formation, is a GCU-specific mRNA interferase in *Escherichia coli*. J. Biol. Chem. *284*, 28746–28753.

Yang, Y.J., Wu, P.J., and Livermore, D.M. (1990). Biochemical characterization of a beta-lactamase that hydrolyzes penems and carbapenems from two *Serratia marcescens* isolates. Antimicrob Agents Chemother *34*, 755–758.

Yarmolinsky, M.B. (1995). Programmed cell death in bacterial populations. Science *267*, 836–837.

Yosef, I., Goren, M.G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. Nucleic Acids Res *40*, 5569–5576.

Yu, Y.T., and Snyder, L. (1994). Translation elongation factor Tu cleaved by a phage-exclusion system. Proc. Natl. Acad. Sci. U.S.A. *91*, 802–806.

Yuan, A.H., and Hochschild, A. (2009). Direct activator/co-activator interaction is essential for bacteriophage T4 middle gene expression. Mol Microbiol *74*, 1018–1030.

Yuan, Y., and Gao, M. (2017). Jumbo Bacteriophages: An Overview. Front Microbiol *8*.

Yuan, A.H., Nickels, B.E., and Hochschild, A. (2009). The bacteriophage T4 AsiA protein contacts the β-flap domain of RNA polymerase. PNAS *106*, 6597–6602.

Zhang, Y., Zhang, J., Hoeflich, K.P., Ikura, M., Qing, G., and Inouye, M. (2003). MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. Mol. Cell *12*, 913–923.

Zhu, Y., Zhang, F., and Huang, Z. (2018). Structural insights into the inactivation of CRISPR-Cas systems by diverse anti-CRISPR proteins. BMC Biol *16*.

Zhuo, S., Clemens, J.C., Stone, R.L., and Dixon, J.E. (1994). Mutational analysis of a Ser/Thr phosphatase. Identification of residues important in phosphoesterase substrate binding and catalysis. J. Biol. Chem. *269*, 26234–26238.

# Chapter 10 Appendices

## 10.1 Goeders et al, 2016

## 10.2 Blower et al, 2017

toxins

MDPI

*Review*

# Structure, Evolution, and Functions of Bacterial Type III Toxin-Antitoxin Systems

**Nathalie Goeders, Ray Chai, Bihe Chen, Andrew Day and George P. C. Salmond ***

Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK; ng394@cam.ac.uk (N.G.); rc636@cam.ac.uk (R.C.); bc407@cam.ac.uk (B.C.); awd33@cam.ac.uk (A.D.)
* Correspondence: gpcs2@cam.ac.uk; Tel.: +44-1223-333-650

Academic Editor: Anton Meinhart
Received: 16 August 2016; Accepted: 19 September 2016; Published: 28 September 2016

**Abstract:** Toxin-antitoxin (TA) systems are small genetic modules that encode a toxin (that targets an essential cellular process) and an antitoxin that neutralises or suppresses the deleterious effect of the toxin. Based on the molecular nature of the toxin and antitoxin components, TA systems are categorised into different types. Type III TA systems, the focus of this review, are composed of a toxic endoribonuclease neutralised by a non-coding RNA antitoxin in a pseudoknotted configuration. Bioinformatic analysis shows that the Type III systems can be classified into subtypes. These TA systems were originally discovered through a phage resistance phenotype arising due to a process akin to an altruistic suicide; the phenomenon of abortive infection. Some Type III TA systems are bifunctional and can stabilise plasmids during vegetative growth and sporulation. Features particular to Type III systems are explored here, emphasising some of the characteristics of the RNA antitoxin and how these may affect the co-evolutionary relationship between toxins and cognate antitoxins in their quaternary structures. Finally, an updated analysis of the distribution and diversity of these systems are presented and discussed.

**Keywords:** abortive infection; altruistic suicide; type III toxin-antitoxin; bacteriophages; quaternary structures; co-evolution; pseudoknotted RNA; endoribonuclease

## 1. Introduction

Toxin-antitoxin (TA) systems are composed of a bacteriostatic or bactericidal toxin and a cognate antidote which is referred to as the antitoxin. In most cases, the antitoxin directly interacts with either the proteinaceous toxin or its mRNA and thus antagonises the deleterious effect of the toxin. In addition to their physical interdependence, they are linked at the genetic level and are often encoded in bicistronic operons with a promoter-proximal antitoxin gene. Save for a few Type I systems, all TA systems are encoded in operons wherein the antitoxin gene is usually found upstream of the toxin-encoding ORF. Nonetheless a few Type II systems such as HigBA and HicAB are exceptions and have a reverse genetic organisation [1]. This transcriptional organisation favours an excess of antitoxin in homeostatic conditions where the toxin is inhibited. The harmful activities of the toxins are due to their interference with essential cellular processes including DNA replication, translation, cell wall synthesis, and maintenance of membrane integrity [2–6].

The active toxin always interacts with its targets as a protein while the nature of the potent antitoxin is either RNA or protein. In addition, antitoxins neutralise their cognate toxins at several levels and act via distinct mechanisms. The nature of the antitoxin and its mode of action underpin the classification of TA systems into five Types (I–V) [7–9]. RNA antitoxins of Type I and III interact with the toxin transcripts (RNA-RNA interactions) or with the toxic protein (RNA-protein interactions) respectively. Most Type I antitoxin RNAs bind the toxin transcript in its 5'UTR region. Formation of this RNA-RNA duplex has two main effects [10]. Firstly, translation of the mRNA into the toxic protein

is hindered as the antitoxin RNA is usually complementary to the region containing the ribosome binding site (RBS) of the toxin transcript or directly competes with ribosomes. In addition to blocking translation initiation, the antitoxin/toxin RNA duplexes are the targets of cellular RNases and thus antitoxin binding ultimately leads to degradation of the toxin transcripts [6]. Aside from Type I TA systems, inhibition of toxin translation is also used by the only Type V system identified so far [7]. In this case, toxin production is turned off directly by the antitoxin which is an RNase that degrades toxin mRNAs and thus directly regulates toxin transcript levels [7]. In Type II, III, and IV systems, translation of the toxin is not directly affected. Type II systems use direct protein-protein interactions where antitoxins either mask the toxin active site or sterically hinder the toxins from reaching their target [1–3]. Type IV antitoxins also prevent toxin-target interactions but achieve this by competing with the toxins for their target, without direct contacts with their cognate toxins [8,9]. Finally, Type III TA systems fall into the category of RNA-protein interactions in which the toxin active site is occluded by the antitoxin RNA [11–14].

TA systems were originally discovered in the late 80s [15,16] and, for a long period, were defined in only two types (Type I and II) which have been extensively studied. In contrast, the archetypal Type III TA system and, subsequently, the Type IV and V systems, were discovered only recently. While data are still comparatively limited for the newer Type IV and V systems, a more holistic image is beginning to emerge regarding Type III systems, with accumulating biochemical, structural, and functional data. This review covers these systems, describing their diversity and toxin-antitoxin/abortive infection bifunctionality and discusses their impact on bacteria-phage co-evolution, given their anti-phage activity.

## 2. Type III TA Systems Are Split into Three Families Which Share the Same Genetic Organisation

The novelty of Type III TA systems involves the nature of RNA-protein interactions between their components [11,12]. These interactions are unusual as the toxin is involved in processing the antitoxin into its active form. More precisely, the antitoxin, a small non-coding RNA (sRNA) composed of several repeats of short nucleotide sequences, is processed into monomers of these repeats by the toxin. The heteromeric complexes adopted by Type III systems during homeostatic conditions are composed of alternating interactions between antitoxin and toxin monomers [12–14]. As for other Types of TA systems, the antitoxin has a shorter half-life than the toxin [11] but the specific details of antitoxin degradation are not yet completely clear.

At the genetic level, Type III TA systems are organised in characteristic bicistronic operons transcribed from a single promoter. The downstream toxin gene is preceded by a Rho-independent terminator that separates it from the upstream short repetitive nucleotide sequences that encode the antitoxin sRNA (Figure 1) [11,17]. Presumably, organisation in operons ensures a higher synthesis of the antitoxin compared to the toxin and thus avoids physiologically precocious, and potentially lethal, toxin activity. Unique to Type III systems toxin expression is further modulated by the presence of the inter-gene Rho-independent terminator. A final regulatory fail-safe may reside in the fact that one antitoxin sRNA will be processed into several monomers that could neutralise two or three toxins, thus further ensuring an appropriately regulated antitoxin:toxin stoichiometry.

Although all Type III TA systems share the same genetic arrangement, they can be further differentiated into three families which are classified according to the amino acid sequence similarities that they share [18]. The subfamilies are called ToxIN, CptIN, and TenpIN where the "I" and "N" represent the antitoxin and toxin components respectively. Thus, for the ToxIN system of *Pectobacterium atrosepticum* the antitoxin is referred to as $ToxI_{Pa}$, the toxin as $ToxN_{Pa}$ and both components as $ToxIN_{Pa}$ [18]. CptIN was named after the *Coprococcus catus* GD/7 system (*Co*Type III Inhibitor/toxiN) and the third family, TenpIN for Type III ENdogenous to *Photorhabdus* Inhibitor/toxIN [18]. While the toxin sequence directly influences the subgroup to which a particular system belongs, it is also interesting to note how their cognate antitoxins differ between and within the subgroups.

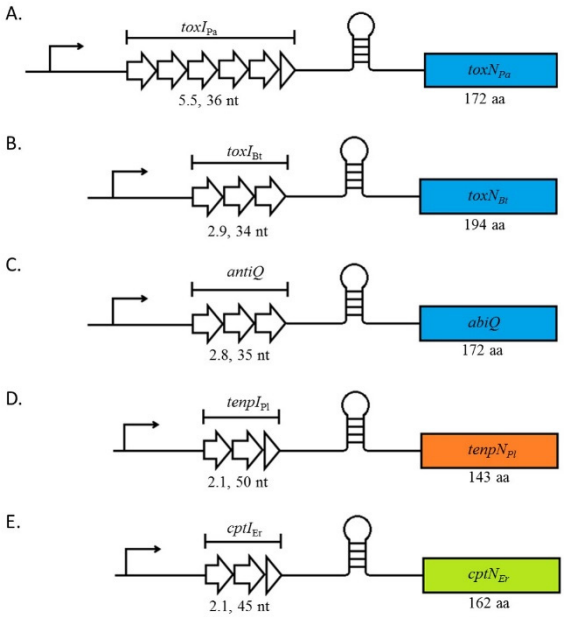



**Figure 1.** Genetic organisation of Type III systems. Type III systems are arranged with the antitoxin gene preceding that of the toxin—separated by a Rho-independent terminator. Five Type III TA loci are shown: (**A**) *toxIN*$_{Pa}$ located on pECA1039 from *Pectobacterium atrosepticum*; (**B**) *toxIN*$_{Bt}$ located on pAW63 from *Bacillus thuringiensis*; (**C**) AbiQ located on pSRQ900 from *Lactococcus lactis*; (**D**) *tenpIN*$_{Pl}$ from the chromosome of *Photorhabdus luminescens*; and finally (**E**) *cptIN*$_{Er}$ from the chromosome of *Eubacterium rectale*.

## 3. Antitoxin Length Is Important for Type III System Functions

Antitoxin repeats are a key feature of Type III systems. The number of repeats varies between systems and they have been shown to be crucial for antitoxin activity. For instance, the antitoxins of the ToxIN$_{Pa}$, ToxIN$_{Bt}$, and AbiQ systems that belong to the ToxIN family, all diverge at the primary sequence level and number of repeats, while the length of the monomers is quite conserved. ToxI$_{Pa}$, ToxI$_{Bt}$, and *antiQ* sRNAs are composed of, respectively, 36 nucleotides repeated 5.5 times, 34 nucleotides repeated 2.9 times, and 35 nucleotides repeated 2.8 times (Figure 1). In vitro, the antitoxin activity can be retained despite increasing or decreasing repeat numbers. However, the range of repeats in which each antitoxin remains functional varies. For instance, 2.5 repeats from 5.5 were necessary and sufficient for ToxI$_{Pa}$ antitoxin to inhibit its toxin [19] while at least 1.8 repeats from 2.8 were essential for the antitoxin activity of *antiQ* [17]. *antiQ* mutants containing 1.8 and 3.8 repeats were readily obtained while clones with only 0.8 of a basic repeat were inviable, suggesting that an incomplete repeat sequence is insufficient to avoid toxicity of AbiQ [17]. In addition to its TA function, the AbiQ system also acts as an abortive infection system against some phages (See below, Section 6.1). This activity is also affected by the number of *antiQ* repeats however the anti-phage activity of the system is altered independently of its toxin neutralising activity. For instance, deletion or addition of one repeat to *antiQ* decreased the phage resistance provided by the AbiQ system, indicating that the length of the wild-type *antiQ* is critical for optimal anti-phage activity. Similarly, mutations in key residues for antitoxin processing led to significant loss of anti-phage activity while a point mutation that affects pseudoknot structure increased anti-phage activity, but did not affect bacterial fitness [17].

## 4. Assembly of the Toxin-Antitoxin Complexes

When the paradigmatic ToxIN$_{Pa}$ system was first discovered, the activity of the toxin component was unknown and mining structural databases with the predicted structure of ToxN$_{Pa}$ gave no meaningful results [11]. Insight into its activity was gained later with the resolution of its crystal

structure and the discovery of the triangular architecture adopted by the three toxin-antitoxin monomers [12]. Resolution of the quaternary structures of further Type III systems showed that this interesting feature of Type III TAs exhibits some variations on a theme where toxin and antitoxin monomers alternate (in hexameric or tetrameric complexes) in which only RNA-protein interactions occur. A hallmark shared by all the structures is that it is the antitoxin processing that leads to the inactive, stable TA complex [12–14]. So far, the core architecture of Type III systems seems to be subfamily specific and likely depends on the length and fold of the antitoxin monomers.

### *4.1. The ToxIN Systems Form Triangular Heterohexamers*

Most of the structural data currently available concerns the ToxIN subfamily. The quaternary structure of the ToxIN$_{Pa}$ and ToxIN$_{Bt}$ systems has been resolved (Figure 2A,B) and bioinformatic analyses predict that the AbiQ system shares the same quaternary architecture [12,13,20]. These crystal structures provided important insights into the mechanism of RNA anti-toxicity.

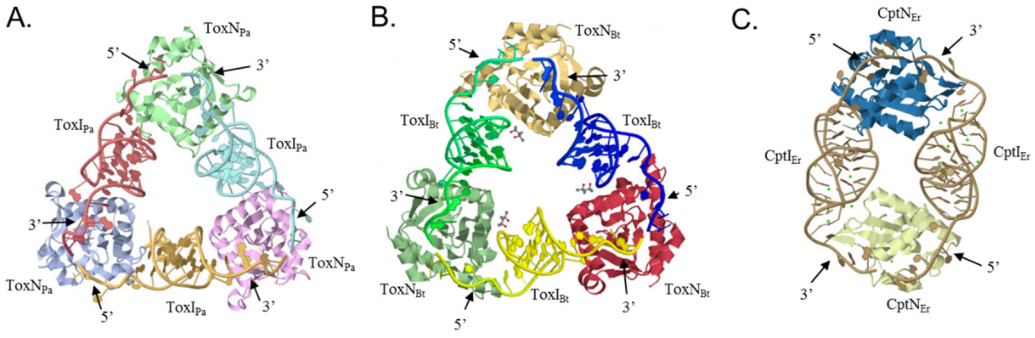



**Figure 2.** Crystal structures of Type III TA systems. (**A**) ToxIN$_{Pa}$ (PDB ID: 2XDD) and (**B**) ToxIN$_{Bt}$ (PDB ID: 4ATO) form heterohexameric complexes [12,13]; (**C**) CptIN$_{Er}$ (PDB ID: 4RMO) assembles into a heterotetrameric complex [14].

Both the ToxIN$_{Pa}$ and ToxIN$_{Bt}$ systems assemble into heterohexameric complexes (Figure 2A,B) that adopt a triangular architecture where toxins occupy the apices and are held together by processed antitoxin monomers [12,13]. In the complexes, each pseudoknotted ToxI RNA is bound head to tail to two separate ToxN monomers and occludes their active sites. Since ToxN$_{Pa}$ and ToxN$_{Bt}$ are both endoribonucleases, cleavage of cognate tandem RNAs into single repeats and assembly into the ToxIN complexes is likely associated with the inhibition of toxicity. Experimental data show that, while ToxN$_{Pa}$ is neutralized by both the processed and precursor forms of ToxI$_{Pa}$, the precursor ToxI$_{Pa}$ is the preferred substrate, indicating the involvement of the antitoxin processing in toxin inhibition [13]. Additionally, the pseudocontinuous arrangement of ToxI units in these structures and the 2′-3′ cyclic phosphates at the 3′ end of the processed antitoxins support this mechanism [12,13]. Finally, ToxIN$_{Pa}$ complexes can self-assemble in vitro from ToxN$_{Pa}$ combined either with processed or precursor ToxI$_{Pa}$ RNAs indicating that the self-assembly of these structures is largely mediated by the antitoxin RNAs and does not require any cellular factors or exogenous energy [13].

Both ToxI$_{Pa}$ and ToxI$_{Bt}$ antitoxin monomers fold into a classic H-type pseudoknot structure flanked by two single-stranded tails. These ends and the adjacent areas of the pseudoknot interact with their respective toxins to stabilise the trimeric structure (Figure 2A,B) [12,13]. Analysis of the crystal structure of ToxIN$_{Pa}$ showed that each ToxI$_{Pa}$ tail interacts with a different ToxN$_{Pa}$ monomer via electropositive grooves where hydrogen bonds occur between the protein and the RNA bases. The antitoxin 3′ end tail containing the 2′-3′ cyclic phosphate is held in place in this groove by five side chains, Tyr29, Lys33, Thr52, Ser53, and Lys55, that form the ToxN$_{Pa}$ active site [12]. Similarly, key areas of ToxI$_{Bt}$ that interact with ToxN$_{Bt}$ are the single stranded tails, where C19 and G20 interact with Lys148 of the toxin, and the 5′ tail of ToxI$_{Bt}$ and U10 interact with a hydrophobic pocket of the ToxN$_{Bt}$ interacting with the 3′ end [13].

### 4.2. CptIN$_{Er}$ Assembles into Heterotetramers

In contrast to the ToxIN systems, CptIN$_{Er}$—currently the only other Type III TA system with a solved crystal structure—assembles into tetramers (Figure 2C). Its quaternary structure is composed of two toxin monomers joined by two antitoxin RNAs [14]. The CptI monomers are longer than the corresponding examples in the ToxI repeats and the fold they adopt probably accounts for the difference in quaternary structures—with the nature of the pseudoknotted RNA perhaps driving the evolution of the toxin-antitoxin complexes.

### 4.3. Type III Antitoxins Form Pseudoknots

All the antitoxins crystallised so far adopt a pseudoknotted fold in their TA complexes. Pseudoknots are a recurrent RNA structural motif in which a loop forms interactions with distal bases outside the loop to form triple stranded structures. Usually the third nucleotide is an adenine enabling A-minor interactions [21]. In the quaternary structures of the TA systems, the antitoxin pseudoknots exhibit three distinct regions that interact with each other through duplex and triplex hydrogen bonds (Figure 3) [12,13]. An important feature is the core of the pseudoknot which contains three internal base triplexes. One of these internal base triplexes (triplex 3, GUU) separates the two based-paired stems of the pseudoknot with interdigitation of a guanine (Figure 3A,B). While key components are conserved in both ToxI pseudoknots, this curious structural aspect was not predicted due to the low sequence homology between the antitoxins. Despite overall similar structures, the RNA-protein interfaces show substantial differences as highlighted by the selective inhibition displayed by both antitoxins and the absence of functional cross-talk between antitoxins and toxins of the *Pectobacterium* and *Bacillus* systems [13].

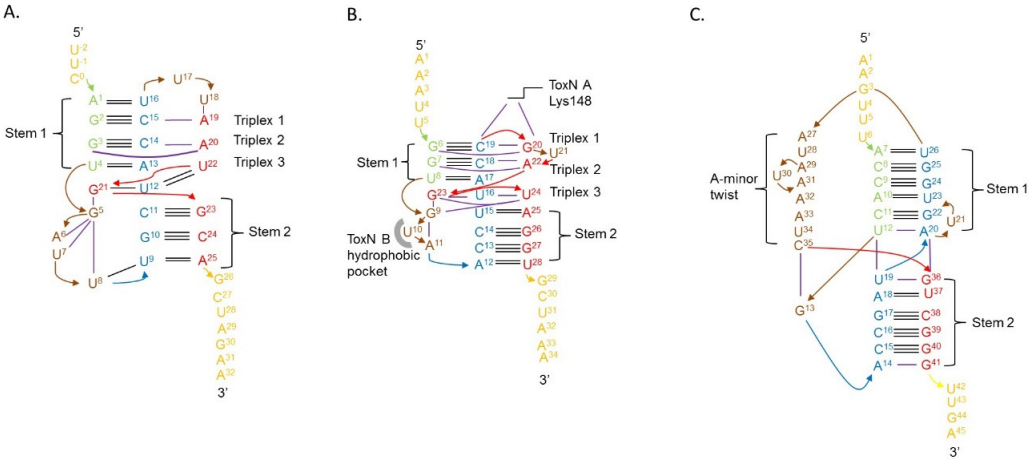



**Figure 3.** Pseudoknot arrangements of Type III antitoxins. (**A**) ToxI$_{Pa}$; (**B**) ToxI$_{Bt}$; and (**C**) CptI$_{Er}$. Base-base hydrogen bonds are shown by black lines. Nucleotides involved in loops are indicated in brown. Corresponding areas for each antitoxin are highlighted by similar colors.

In contrast, CptI$_{Er}$ monomers fold into an H-type pseudoknot distinctly different from the fold of the antitoxins of the ToxIN family. Firstly, CptI$_{Er}$ has two coaxial stems assembled entirely from duplex base pairing (Figure 3C) as opposed to the triplex base pairing seen in the other two antitoxins [14]. Even among RNA pseudoknots, the loops of the CptI$_{Er}$ pseudoknots are unusual. Indeed, loop 1 (L1) interacts with stem 2 (S2) while loop 2 (L2) interacts with stem 1 (S1), a configuration that leads often to triplex base pairing [21]. These canonical pseudoknot interactions are found in the ToxI RNAs where two triplex base pairs occur from the interaction of L2 with S1, although there is no triplex pairing with L1 and S2 (Figure 3A,B) [12,13]. The CptI$_{Er}$ pseudoknot is more atypical as L1 is extremely short and only consists of a single base that does not interact with S2 and is instead held in place by interactions with the 3′ end of L2 (Figure 3C) [14]. L2 of CptI$_{Er}$ is adenine rich and longer than in the

other antitoxins and forms a novel counter-clockwise A-minor twist motif [14]. This conformation appears to be necessary for its antitoxin activity, as suggested by experiments on substitution and deletion mutations that disrupted key features of this twist [14]. As with the ToxI-type antitoxins, functional cross inhibition experiments have confirmed that CptI antitoxins also show high toxin specificity [14].

## 5. Type III Toxins Share a Common Fold and Activity

To date, the structures of four Type III toxins have been solved; three of which belong to members of the ToxIN family (ToxIN$_{Pa}$, ToxIN$_{Bt}$ and AbiQ) and, more recently, a member of the CptIN family. All toxins share a globular β core surrounded by α-helices and loops. While the core structure is conserved, most variations are found on the surface. These variations are thought to account for cleavage and antitoxin specificity of the toxins. All Type III toxins tested so far for their mechanistic action have been shown to be endoRNases that cleave mRNAs in adenine-rich regions but with slightly different substrate specificities: ToxN$_{Pa}$ AA/AU, ToxN$_{Bt}$ A/AAAA, and AbiQ A/AAA [13,20]. As for other toxins that inhibit translation, Type III toxins initially have a bacteriostatic effect on growth [3,11,20] that ultimately leads to lethality.

### *5.1. The ToxN Family*

ToxN$_{Pa}$ and ToxN$_{Bt}$ have the same highly twisted core of six anti-parallel β-sheets and five main variable loop regions (Figure 4A,B) [12,13]. Four of these loops, especially the long kinked helix H3, act as the major sites of interaction between the toxins and their respective antitoxins while the fifth encompasses the active site. In line with the structural data, bioinformatics analysis of the sequences of the other members of the ToxN family showed that the core fold is conserved while most sequence variability clusters in regions corresponding to these five loops [12]. The structure of AbiQ also encompasses a core of six-stranded anti-parallel β-sheets surrounded by 6 α-helices (Figure 4C) [20]. Surprisingly, the RNase activity of AbiQ was only eliminated by one site-directed mutation in Ser51Leu which is thought to make the nucleophilic attack during the cleavage [20]. The RNase activity of AbiQ was retained when replacing the Ser51 by a threonine, the equivalent residue found in the active site of ToxN [12,13,20].

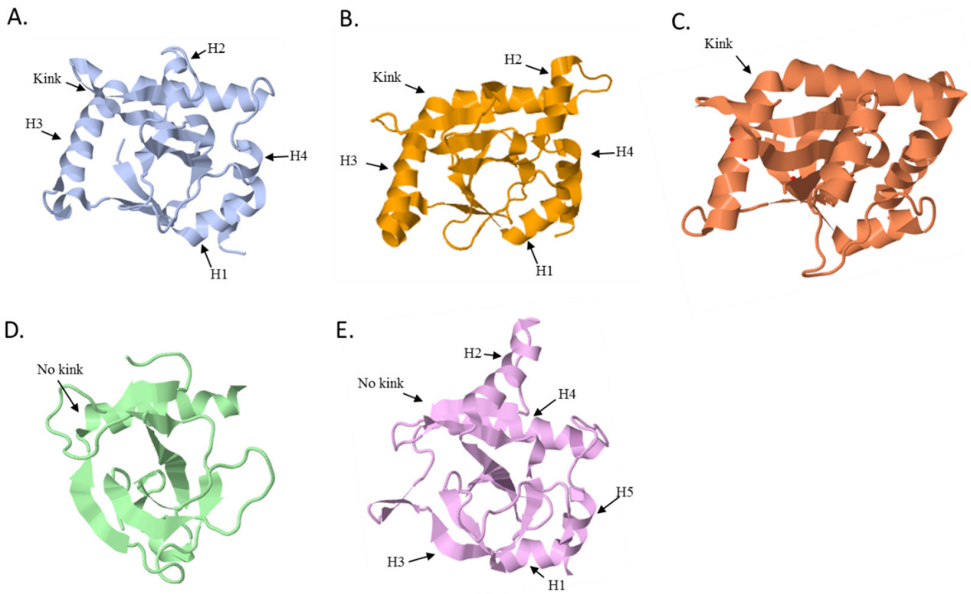



**Figure 4.** Structures of the toxins. (**A**) Structure of ToxN$_{Pa}$ (PDB ID: 2XDD, [12]); (**B**) ToxN$_{Bt}$ (PDB ID: 4ATO, [13]); (**C**) AbiQ (PDB ID: 4GLK, [20]); (**D**) Kid (PDB ID: 1M1F, [22]); and (**E**) CptN$_{Er}$ (PDB ID: 4RMO, [14]). The presence or absence of the kink in helix 3 (H3) is indicated for each toxin.

### *5.2. The CptN Family*

$CptN_{Er}$ shares the same core fold found in the other toxins despite its lower primary sequence identity with members of the ToxN family (Figure 4E). In $CptN_{Er}$, the highly twisted anti-parallel β-sheets that form the core are surrounded by four α-helices that make extensive interactions with the cognate antitoxin sRNA [14]. Compared to the ToxN family, in which $ToxN_{Pa}$, $ToxN_{Bt}$, and AbiQ have a kink in helix three [12,13,20] the equivalent helix in $CptN_{Er}$, helix H4, is much shorter [14]. However many of the hydrogen bonding networks in the RNase active site are also conserved suggesting that this toxin is also an endoRNase.

### *5.3. Type III Toxins Share Homology with Type II Toxins*

Unexpectedly, the structure of Type III toxins shows significant homology with toxins from the Type II MazF/Kid/CcdB family (Figure 4). Members of this Type II family share a similar fold despite low sequence similarity and different mechanisms of action. For instance, Kid and MazF toxins inhibit translation by acting as endoRNases while CcdB toxins interact with DNA gyrase and affect DNA replication [23–25]. Type II and Type III toxins share very low sequence similarity, e.g., only 11% identity between $ToxN_{Pa}$ and Kid, but their respective structures encompass a similar β core fold region surrounded by helices and loops. Interestingly, despite sharing the same overall structure and the same molecular activity i.e., endoRNases, the active sites of ToxN and Kid do not match well [11,25]. Less surprisingly, differences are also found in the regions that interact with their respective antitoxins. For instance, the helix 3 region of $ToxN_{Pa}$ overlays with the Kid structure but is greatly extended in the N-terminal section of the $ToxN_{Pa}$, the main site of ToxI recognition. The equivalent helix in $CptN_{Er}$, helix H4, is much shorter and, as such, resembles more closely to the Kid toxin (Figure 4A,D,E).

## 6. Functions of Type III TA System

Two functions have been ascribed to Type III TA systems so far. Historically, the first function was their abortive infection (Abi) activity that protects bacterial populations from invading bacteriophages. Their second known function involves their role in stabilisation of plasmids.

### *6.1. Abortive Infection*

Some Type III systems are active protagonists in the "molecular arms race" that occurs between bacteria and phages. As these outnumber bacteria 10-fold in some environments, the prokaryotic hosts are potentially subject to constant viral predation [26]. In response to this acute selective pressure on microbial communities, bacteria have evolved a large array of defense mechanisms to circumvent the lethal impact of viral infections [27,28].

At least some of these Type III systems are bifunctional and can act as abortive infection (Abi) systems in addition to their TA functionality. This bifunctionality is possible as both activities require the involvement of a self-poisoning protein. Abi mechanisms rely on a toxic protein that can be viewed, from an evolutionary perspective, as inducing the altruistic suicide of infected cells to prevent phage propagation in bacterial populations. Even though 23 different Abi systems have been reported, mostly carried on lactococcal plasmids, there is still a paucity of information on the molecular mechanisms involved in their function, with the exception of TA/Abi systems for which some insight has been acquired recently [29]. Abi systems have been shown to act at different steps of the phage replication cycle i.e., from DNA replication to the lysis step—but ultimately they all lead to the death of the infected bacteria [29]. TA/Abi bifunctionality is not restricted to Type III systems as TA systems from Type I to IV have been shown to protect bacteria from bacteriophages [8,11,30–32]. While molecular details are still elusive in most cases, a working model is that—following toxin activation by the infecting phage—dissemination of phage progeny in the bacterial population is restricted. Toxin activation could rely on the differential stability of the antitoxin and toxin products.

Post-infection, this characteristic of TA systems could favour a state of free toxins upon the depletion of the shorter lived antitoxins.

The ToxIN$_{Pa}$ system of *P. atrosepticum* was identified originally through its phage resistance capacity encoded by a cryptic plasmid. Subcloning of different plasmid fragments eventually narrowed the anti-phage activity down to the ToxIN system genetic module [11]. ToxIN$_{Pa}$ can inhibit a spectrum of phages infecting multiple Gram-negative enterobacteria such as *P. atrosepticum*, *Escherichia coli*, and *Serratia* spp. ([10] and unpublished data). This dual Abi and TA functionality was shown later to be shared by some other Type III TA systems. The chromosomal *tenpIN* locus from *Photorhabdus luminescens*, effectively aborts environmental coliphages when expressed from a plasmid in *E. coli* [11,18,33] and, likewise, AbiQ is effective, in Gram-positive *Lactococcus lactis*, against members of the common 936 and c2 phage groups as well as rarer lactococcal phages. When expressed in *E. coli*, AbiQ also protects this Gram-negative host from some coliphages [33]. Therefore, the protective effect against phages appears to be independent of the organism in which the systems are present [11,18,33]. Phage resistance may be also independent of the original genetic location as both plasmid and chromosomal systems can effectively inhibit phage replication [11,18]. Similarly, no correlation between the Type III TA systems families and Abi activity can be drawn currently. For instance, despite proficient anti-phage activity of the ToxIN$_{Pa}$ and AbiQ systems, no anti-phage activity could be shown for the closely related ToxIN$_{Bt}$ system in its native host [13] and while both phenotypes have been observed for members of the ToxIN family, no anti-phage activity has yet been observed for the two tested CptIN systems from *Eubacterium rectale* and *Ruminococcus torque* [14,18].

A further aspect of anti-phage activity of Type III systems is the specificity these systems show against subgroups of sensitive phages [11,18]. No correlation has been found between phage families (*Myoviridae*, *Siphoviridae*, and *Podoviridae*) and abortive infection [33,34]. It is not known whether phages that are not aborted by the TA/Abi systems have naturally and actively evolved resistance mechanisms against these systems in the perpetual arms race that occurs between them and their bacterial hosts, or whether they "simply" do not activate the toxin of these systems. Some examples of mechanisms selected by phages to avoid Abi systems in general (as well as TA/Abi systems) are known in the literature [34]. For instance, point mutations in key phage products have been identified as the basis of escape mechanisms. More recently it was shown that phage T4 has an ADP ribosyltransferase that chemically modifies the Type II MazF toxin to downregulate its activity and thus avoid this TA/Abi system [32]. Another phage escape mechanism includes two examples where the phages evolved mimics of the toxin's substrates. In the case of the Type II RnlAB/LsoAB systems of *E. coli* and phage T4, the phage uses the Dmd protein which acts like a Type II antitoxin and directly interacts with the toxins. In contrast to the canonical antitoxins of RnlA and LsoA, Dmd is able to cross-neutralise several toxins. Recent crystallographic studies showed that Dmd is thought to have a different inhibition mechanism and directly interacts with the toxin active site by mimicking the toxin substrates [35].

In the Type III systems, the ToxIN-sensitive phage TE was shown to produce low frequency spontaneous mutants that escaped the ToxIN$_{Pa}$ system [36]. Analysis of the escape mutants revealed that the majority of the mutants had extended a short viral sequence similar to the repeats of the ToxI sRNA into a 'pseudo-ToxI' which functionally suppressed the toxin. In one case, recombination had allowed the phage escape mutant to obtain natural ToxI repeats from the original plasmid antitoxin sequences. Both scenarios allowed phage replication unaffected by the Abi/TA system [36].

The precise molecular mechanisms of these bifunctional TA/Abi systems are still under investigation. In agreement with the current model, the toxin endoRNase activity and the Abi phenotype have been shown to be linked [19]. However, this may not be a universal situation because mutagenesis of key toxin residues can lead to the loss of the Abi phenotype despite retained RNase activity [20] indicating that the details of the toxin activation are more subtle. This leads to the question as to how Type III TA/Abi systems are actually activated and the molecular basis of the differential sensitivity of different phages. Phage products that are involved in the activation of Type III

TA/Abi systems might directly interact with the toxin, the antitoxin, or both, to prevent assembly of the TA heterocomplexes, or to disrupt them. Alternatively, a phage product might interrupt transcription of the TA operon or affect antitoxin RNA stability, thus causing imbalance in the TA components. Both scenarios would lead to an excess of free toxin, which could inhibit mRNA translation, leading to cell growth arrest and thereby ultimately prevent the release of phage progeny. The nature of the phage products involved and the way they interact (directly or indirectly) with the TA system is unknown. Experimental data for the ToxIN$_{Pa}$ and the AbiQ systems showed that toxin levels are not affected during phage infections nor is transcription of the TA systems increased [19,20]. However transcripts of the AbiQ system, expressed constitutively before infection, decrease during the infection, possibly as part of a general infection phenomenon [20]. Based on experimental data, it is thought that the TA systems are activated at late steps of phage replication, at least in the AbiQ system. It has been shown that phage DNA is replicated in infected cells in the presence of the AbiQ system, as indicated by the accumulation of concatemeric viral DNA [37].

### *6.2. Plasmid Inheritance through Addiction*

Another function known for Type III systems is plasmid addiction. In contrast to the Abi phenomenon where TA systems protect bacteria from invading DNA, when acting as "addiction" modules they ensure the stable inheritance of plasmids in bacterial populations.

This feature of TA systems also relies on the antagonism of the self-poisoning essence of the toxin and the labile nature of the antitoxin [11]. It supposes a continuous synthesis of the antitoxin in order to keep the toxin in an inhibited state and thus renders the cell 'addicted' to the TA system. Consequently, when present on mobile genetic elements such as plasmids, these "addiction" modules promote the maintenance of the DNA molecule encoding the TA system and ensure the continued presence of plasmids in bacterial populations via killing of plasmid-free segregant cells by the toxins [15,38]. Historically, this mechanism has also been named post-segregational killing (PSK) as lethality arises when plasmids are not segregated into both daughter cells during cell division [38]. Alternatively, given the initially bacteriostatic—rather than bactericidal—nature of Type III toxins [11,20], mis-segregation could lead to a transitory growth inhibition. Bacteria that retained the plasmid would then be able to outgrow those that "lost" it, thus ensuring plasmid maintenance at the population level.

Plasmid addiction is a function found in many TA systems of Type I and II and has also been shown for some Type III TA systems. For example, ToxIN$_{Pa}$ and CptIN$_{Er}$ increased plasmid retention in *E. coli* W3110 to 100%, compared to 50% loss of the control vector [13,14]. Notably, the fact that CptIN$_{Er}$ is located on the chromosome of the original host, yet could promote plasmid maintenance when cloned, suggests that CptIN$_{Er}$ can promote its own retention and might be disseminated through horizontal gene transfer [14].

In addition to the post-segregational killing during the exponential growth phase, one Type III system has been show to ensure plasmid maintenance in stress conditions leading to sporulation of *Bacillus subtilis*. During sporulation, bacteria differentiate into spores, a metabolically dormant cell-type that can survive adverse environments and switch back to vegetative growth in favorable conditions. The frequency of plasmid loss during sporulation is higher than in vegetative growth [39] and the rate of plasmid loss can vary from 5% to 95% in *Bacillus* [40]. The *B. subtilis* ToxIN$_{Bt}$ system favours plasmid retention in this specific context by decreasing plasmid loss around 10-fold (from 58% for the control plasmid to 6% when the TA system was present) [39]. This was achieved by reducing the proportion of cells that form mature spores—probably by the action of the toxin in plasmid-free forespores. Only forespores that inherited a plasmid copy were able to mature and so this mechanism ensures plasmid maintenance throughout the environmental stress that precipitated sporulation. It is formally possible that other types of TA systems might affect plasmid retention by similar routes during sporulation [39].

## 7. New Bioinformatic Analysis Reveals a Significant Increase in Potential Novel Systems

Mining databases for ORFs and subsequent analysis of the proteins encoded by them can be a powerful tool to gain evolutionary insight into genes encoding TA systems [41,42]. Investigating the occurrence of Type III systems was initially complicated as bioinformatic analysis of non-coding RNAs is still problematic. Solving the structure of ToxN made it possible to identify new members of the Type III TA family by performing de novo structure-based homology searches [18]. A set of criteria was used to determine which ToxN structural homologues found by FUGUE [43], a program for sequence-structure comparison, are truly from Type III TA systems. The logic of the definition was as follows: the ToxN homologues should be preceded by a palindromic repeat acting as a Rho-independent terminator, which in turn should be preceded by sequences with a similar organization to *toxI*—composed of a tandem array of nucleotide repeats. Subsequently, samples from each family were taken and exhaustive BLASTp searches were performed, yielding a final list of 125 putative type III TA systems. The majority of hits were found on bacterial chromosomes and plasmids of Firmicutes, Fusobacteria, and Proteobacteria as well as Archaea. One *toxIN* locus was also encoded on a prophage [18]. Using these criteria, 37 putative Type III loci were identified and were divided into three independent families according to their protein sequences as described previously [18].

Given the near exponential increase of sequenced genomes available in databases during the last few years, we did a new BLASTp search. Using the amino acid sequences of the characterized ToxN$_{Pa}$, ToxN$_{Bt}$, AbiQ, TenpN$_{Pl}$, and CptN$_{Er}$ toxins as input, we did an initial BLASTp search against the Bacteria and Viruses databases of Uniprot. We identified 603 potential Type III toxins with an *E*-value <0.001. A selection of these predicted toxins were further analysed following the previously established criteria. After validation as members of putative Type III TA systems, these toxins were used for a further round of BLASTp. The new round of BLASTp searches with the selected toxin homologues gave a few new putative toxins indicating that the search is exhaustive. The different toxin groups seem to cluster to particular phylogenetic groups. For instance, the ToxN family is primarily found in Firmicutes, Fusobacteria and occasionally other species. Systems from this group also seem associated with the mobilome as they can be found on plasmid and chromosomal locations, sometimes in the close vicinity of transposase elements. Despite the increased number of predicted systems, Type III TA systems appear largely restricted to a limited group of phyla as only 8 out of the currently estimated 74–76 bacterial phyla have at least one potential system [44]. Whether Type III TA systems are indeed only confined to a restricted number of bacteria, or whether the current distribution pattern is biased by the limited phylogenetic breadth of genome sequencing, is unknown. It is intriguing that Type III TA systems seem to be confined to a few phylogenetic groups. As the toxin mode of action, i.e., cleaving RNAs, is a general one that does not require any known external factors, and given that they are found on plasmids, one might expect them to be spread through more phylogenetic groups, in a similar way to the Type II systems where the majority of toxins are also RNases [41,42,45].

Advances in technologies such as single-cell genomics and metagenomics will enable the sequencing of uncultivated organisms from diverse habitats [44,46] and might affect the picture of TA system distribution as the current "tree of life" expands and culturing-associated biases are overcome [43,47]. When placing the consensus antitoxin repeats next to a phylogenetic tree based on the protein toxin sequences, it becomes clear that the antitoxins of closely related toxins are conserved as well (Figure 5). This consistency provided evidence for the notion of co-evolution of toxin and antitoxin components. The length and the primary sequence of the consensus repeats are usually quite conserved within the different toxin groups. On the other hand, the number of repeats is prone to slight variations.

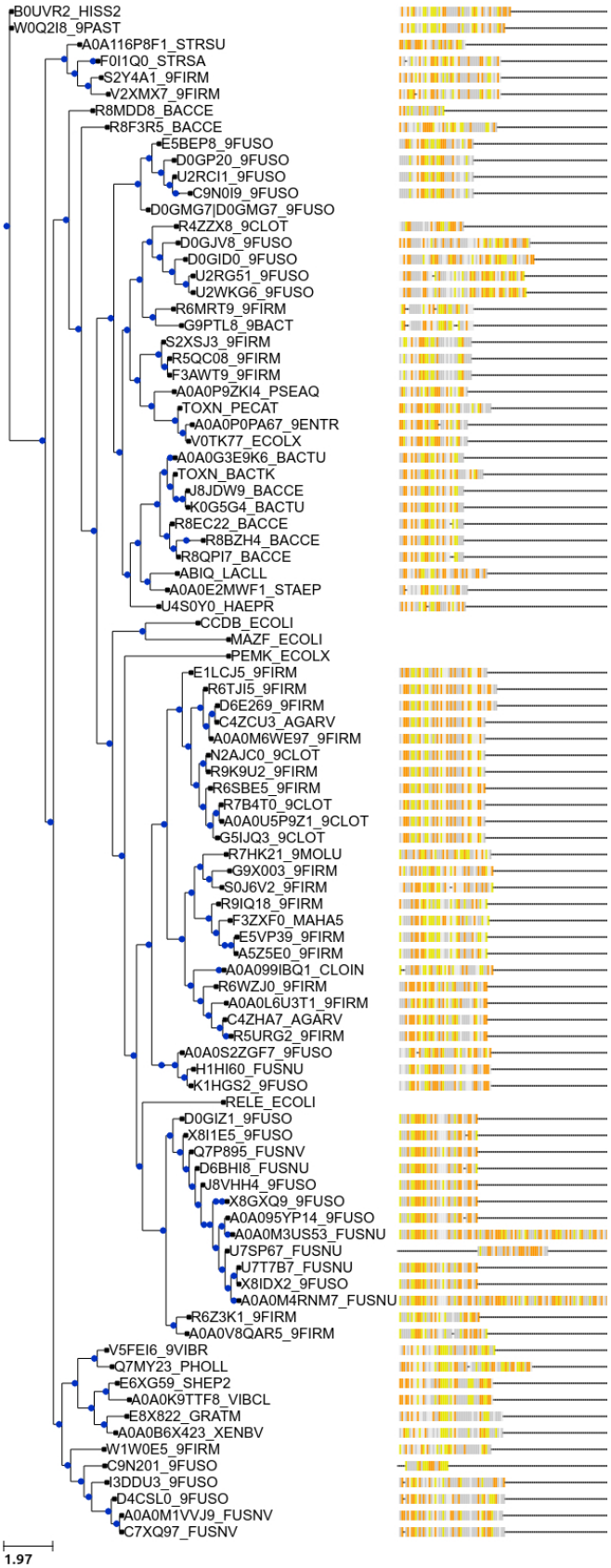



**Figure 5.** Phylogenetic tree based on representative toxins of putative Type III TA systems associated with the consensus repeats of the antitoxins: adenine (grey), guanine (white), cytosine (yellow), and uracil (orange). Toxin protein sequences were aligned using MUSCLE [48]. The tree has been constructed using PhyML and nearest neighbour interchange [49,50] and the figure has been made using Ete3 [51].

## 8. Conclusions

Work in recent years has increased our understanding of bacterial Type III TA systems. However, it is abundantly clear that our understanding of the molecular mechanisms involved in Type III systems and their activation is still rudimentary. These TA systems provide valuable reagents for fundamental biochemical studies on protein:RNA interactions and the study of quaternary nucleoprotein complex assemblies. The ecological and fitness implications of bacterial carriage of Type III TA systems have yet to be investigated. Furthermore, the role of Type III systems in the evolution, replication, and physiology of both bacteria and their viral parasites warrants deeper consideration.

**Acknowledgments:** Work in the Salmond lab is supported by the BBSRC, UK; N.G. was supported by the Fonds National de la Recherche Luxembourg (9118191); B.C. was supported by a Cambridge International Scholarship from the Cambridge Commonwealth, European & International Trust; and A.D. was supported by a BBSRC -DTP studentship.

**Author Contributions:** All authors were involved in the writing and revision of the review article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Jurenaite, M.; Markuckas, A.; Sužiedeliene, E. Identification and characterization of type II toxin-antitoxin systems in the opportunistic pathogen *Acinetobacter baumannii*. *J. Bacteriol.* **2013**, *195*, 3165–3172. [CrossRef] [PubMed]
2. Yamaguchi, Y.; Inouye, M. An endogenous protein inhibitor, YjhX (TopAI), for topoisomerase I from *Escherichia coli*. *Nucleic Acids Res.* **2015**, *43*, 10387–10396. [PubMed]
3. Goeders, N.; Van Melderen, L. Toxin-antitoxin systems as multilevel interaction systems. *Toxins* **2014**, *6*, 304–324. [CrossRef] [PubMed]
4. Harms, A.; Stanger, F.V.; Scheu, P.D.; de Jong, I.G.; Goepfert, A.; Glatter, T.; Gerdes, K.; Schirmer, T.; Dehio, C. Adenylylation of gyrase and topo IV by FicT toxins disrupts bacterial DNA topology. *Cell Rep.* **2015**, *12*, 1497–1507. [CrossRef] [PubMed]
5. Sofos, N.; Xu, K.; Dedic, E.; Brodersen, D.E. Cut to the chase—Regulating translation through RNA cleavage. *Biochimie* **2015**, *114*, 10–17. [CrossRef] [PubMed]
6. Brielle, R.; Pinel-Marie, M.-L.; Felden, B. Linking bacterial type I toxins with their actions. *Curr. Opin. Microbiol.* **2016**, *30*, 114–121. [CrossRef] [PubMed]
7. Wang, X.; Lord, D.M.; Cheng, H.-Y.; Osbourne, D.O.; Hong, S.H.; Sanchez-Torres, V.; Quiroga, C.; Zheng, K.; Herrmann, T.; Peti, W.; et al. A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. *Nat. Chem. Biol.* **2012**, *8*, 855–861. [CrossRef] [PubMed]
8. Dy, R.L.; Przybilski, R.; Semeijn, K.; Salmond, G.P.C.; Fineran, P.C. A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism. *Nucleic Acids Res.* **2014**, *42*, 4590–4605. [CrossRef] [PubMed]
9. Masuda, H.; Tan, Q.; Awano, N.; Yamaguchi, Y.; Inouye, M. A novel membrane-bound toxin for cell division, CptA (YgfX), inhibits polymerization of cytoskeleton proteins, FtsZ and MreB, in *Escherichia coli*. *FEMS Microbiol. Lett.* **2012**, *328*, 174–181. [CrossRef] [PubMed]
10. Wen, J.; Fozo, E. sRNA antitoxins: More than one way to repress a toxin. *Toxins* **2014**, *6*, 2310–2335. [CrossRef] [PubMed]
11. Fineran, P.C.; Blower, T.R.; Foulds, I.J.; Humphreys, D.P.; Lilley, K.S.; Salmond, G.P. The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 894–899. [CrossRef] [PubMed]
12. Blower, T.R.; Pei, X.Y.; Short, F.L.; Fineran, P.C.; Humphreys, D.P.; Luisi, B.F.; Salmond, G.P.C. A processed noncoding RNA regulates an altruistic bacterial antiviral system. *Nat. Struct. Mol. Biol.* **2011**, *18*, 185–190. [CrossRef] [PubMed]
13. Short, F.L.; Pei, X.Y.; Blower, T.R.; Ong, S.-L.; Fineran, P.C.; Luisi, B.F.; Salmond, G.P.C. Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E241–E249. [CrossRef] [PubMed]

14. Rao, F.; Short, F.L.; Voss, J.E.; Blower, T.R.; Orme, A.L.; Whittaker, T.E.; Luisi, B.F.; Salmond, G.P.C. Co-evolution of quaternary organization and novel RNA tertiary interactions revealed in the crystal structure of a bacterial protein-RNA toxin-antitoxin system. *Nucleic Acids Res.* **2015**, *43*, 9529–9540. [CrossRef] [PubMed]
15. Ogura, T.; Hiraga, S. Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 4784–4788. [CrossRef] [PubMed]
16. Gerdes, K.; Rasmussen, P.B.; Molin, S. Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 3116–3120. [CrossRef] [PubMed]
17. Bélanger, M.; Moineau, S. Mutational analysis of the antitoxin in the lactococcal type III toxin-antitoxin system AbiQ. *Appl. Environ. Microbiol.* **2015**, *81*, 3848–3855. [CrossRef] [PubMed]
18. Blower, T.R.; Short, F.L.; Rao, F.; Mizuguchi, K.; Pei, X.Y.; Fineran, P.C.; Luisi, B.F.; Salmond, G.P.C. Identification and classification of bacterial Type III toxin-antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Res.* **2012**, *40*, 6158–6173. [CrossRef] [PubMed]
19. Blower, T.R.; Fineran, P.C.; Johnson, M.J.; Toth, I.K.; Humphreys, D.P.; Salmond, G.P.C. Mutagenesis and functional characterization of the RNA and protein components of the *toxIN* abortive infection and toxin-antitoxin locus of *Erwinia*. *J. Bacteriol.* **2009**, *191*, 6029–6039. [CrossRef] [PubMed]
20. Samson, J.E.; Spinelli, S.; Cambillau, C.; Moineau, S. Structure and activity of AbiQ, a lactococcal endoribonuclease belonging to the type III toxin-antitoxin system: AbiQ, a type III toxin-antitoxin system. *Mol. Microbiol.* **2013**, *87*, 756–768. [CrossRef] [PubMed]
21. Butcher, S.E.; Pyle, A.M. The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc. Chem. Res.* **2011**, *44*, 1302–1311. [CrossRef] [PubMed]
22. Hargreaves, D.; Santos-Sierra, S.; Giraldo, R.; Sabariegos-Jareño, R.; de la Cueva-Méndez, G.; Boelens, R.; Díaz-Orejas, R.; Rafferty, J.B. Structural and functional analysis of the Kid toxin protein from *E. coli* plasmid R1. *Structure* **2002**, *10*, 1425–1433. [CrossRef]
23. Loris, R.; Dao-Thi, M.-H.; Bahassi, E.M.; Van Melderen, L.; Poortmans, F.; Liddington, R.; Couturier, M.; Wyns, L. Crystal structure of CcdB, a topoisomerase poison from *E. coli*. *J. Mol. Biol.* **1999**, *285*, 1667–1677. [CrossRef] [PubMed]
24. Kamada, K.; Hanaoka, F.; Burley, S.K. Crystal structure of the MazE/MazF complex: Molecular bases of antidote-toxin recognition. *Mol. Cell* **2003**, *11*, 875–884. [CrossRef]
25. Diago-Navarro, E.; Hernandez-Arriaga, A.M.; López-Villarejo, J.; Muñoz-Gómez, A.J.; Kamphuis, M.B.; Boelens, R.; Lemonnier, M.; Díaz-Orejas, R. *parD* toxin-antitoxin system of plasmid R1—Basic contributions, biotechnological applications and relationships with closely-related toxin-antitoxin systems. *FEBS J.* **2010**, *15*, 3097–3117. [CrossRef] [PubMed]
26. Weinbauer, M.G. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* **2004**, *28*, 127–181. [CrossRef] [PubMed]
27. Labrie, S.J.; Samson, J.E.; Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **2010**, *8*, 317–327. [CrossRef] [PubMed]
28. Koskella, B.; Brockhurst, M.A. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **2014**, *38*, 916–931. [CrossRef] [PubMed]
29. Chopin, M.-C.; Chopin, A.; Bidnenko, E. Phage abortive infection in lactococci: Variations on a theme. *Curr. Opin. Microbiol.* **2005**, *8*, 473–479. [CrossRef] [PubMed]
30. Pecota, D.C.; Wood, T.K. Exclusion of T4 phage by the *hok/sok* killer locus from plasmid R1. *J. Bacteriol.* **1996**, *178*, 2044–2050. [PubMed]
31. Hazan, R.; Engelberg-Kulka, H. *Escherichia coli mazEF*-mediated cell death as a defense mechanism that inhibits the spread of phage P1. *Mol. Genet. Genom.* **2004**, *272*, 227–234. [CrossRef] [PubMed]
32. Alawneh, A.M.; Qi, D.; Yonesaki, T.; Otsuka, Y. An ADP-ribosyltransferase Alt of bacteriophage T4 negatively regulates the *Escherichia coli* MazF toxin of a toxin-antitoxin module: ADP-ribosylation of *E. coli* MazF by T4 Alt. *Mol. Microbiol.* **2016**, *99*, 188–198. [CrossRef] [PubMed]
33. Samson, J.E.; Belanger, M.; Moineau, S. Effect of the abortive infection mechanism and type III toxin/antitoxin system AbiQ on the lytic cycle of *Lactococcus lactis* Phages. *J. Bacteriol.* **2013**, *195*, 3947–3956. [CrossRef] [PubMed]

34. Blower, T.R.; Evans, T.J.; Przybilski, R.; Fineran, P.C.; Salmond, G.P.C. Viral evasion of a bacterial suicide system by RNA–based molecular mimicry enables infectious altruism. *PLoS Genet.* **2012**, *8*, e1003023. [CrossRef] [PubMed]
35. Samson, J.E.; Magadan, A.H.; Sabrie, M.; Moineau, S. Revenge of the phages: Defeating bacterial defences. *Nat. Rev. Microbiol.* **2013**, *11*, 675–687. [CrossRef] [PubMed]
36. Wei, Y.; Gao, Z.; Zhang, H.; Dong, Y. Structural characterizations of phage antitoxin Dmd and its interactions with bacterial toxin RnlA. *Biochem. Biophys. Res. Commun.* **2016**, *472*, 592–597. [CrossRef] [PubMed]
37. Emond, E.; Dion, E.; Walker, S.A.; Vedamuthu, E.R.; Kondo, J.K.; Moineau, S. AbiQ, an abortive infection mechanism from *Lactococcus lactis*. *Appl. Environ. Microbiol.* **1998**, *64*, 4748–4756. [PubMed]
38. Sengupta, M.; Austin, S. Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. *Infect. Immun.* **2011**, *79*, 2502–2509. [CrossRef] [PubMed]
39. Short, F.L.; Monson, R.E.; Salmond, G.P.C. A Type III protein-RNA toxin-antitoxin system from *Bacillus thuringiensis* promotes plasmid retention during spore development. *RNA Biol.* **2015**, *12*, 933–937. [CrossRef] [PubMed]
40. Turgeon, N.; Laflamme, C.; Ho, J.; Duchaine, C. Evaluation of the plasmid copy number in *B. cereus* spores, during germination, bacterial growth and sporulation using real-time PCR. *Plasmid* **2008**, *60*, 118–124. [CrossRef] [PubMed]
41. Makarova, K.S.; Wolf, Y.I.; Koonin, E.V. Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol. Direct* **2009**, *4*, 19. [CrossRef] [PubMed]
42. Leplae, R.; Geeraerts, D.; Hallez, R.; Guglielmini, J.; Dreze, P.; Van Melderen, L. Diversity of bacterial type II toxin-antitoxin systems: A comprehensive search and functional analysis of novel families. *Nucleic Acids Res.* **2011**, *39*, 5513–5525. [CrossRef] [PubMed]
43. Shi, J.; Blundell, T.L.; Mizuguchi, K. FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol.* **2001**, *310*, 243–257. [CrossRef] [PubMed]
44. Hug, L.A.; Baker, B.J.; Anantharaman, K.; Brown, C.T.; Probst, A.J.; Castelle, C.J.; Butterfield, C.N.; Hernsdorf, A.W.; Amano, Y.; Ise, K.; et al. A new view of the tree of life. *Nat. Microbiol.* **2016**, *1*, 16048. [CrossRef] [PubMed]
45. Guglielmini, J.; Van Melderen, L. Bacterial toxin-antitoxin systems: Translation inhibitors everywhere. *Mob. Genet. Elem.* **2011**, *1*, 283–306. [CrossRef] [PubMed]
46. Rinke, C.; Schwientek, P.; Sczyrba, A.; Ivanova, N.N.; Anderson, I.J.; Cheng, J.-F.; Darling, A.; Malfatti, S.; Swan, B.K.; Gies, E.A.; et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **2013**, *499*, 431–437. [CrossRef] [PubMed]
47. Spang, A.; Ettema, T.J.G. Microbial diversity: The tree of life comes of age. *Nat. Microbiol.* **2016**, *1*, 16056. [CrossRef] [PubMed]
48. Eddy, S.R. What is a hidden Markov model? *Nat. Biotechnol.* **2004**, *22*, 1315–1316. [CrossRef] [PubMed]
49. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [CrossRef] [PubMed]
50. Guindon, S.; Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **2003**, *52*, 696–704. [CrossRef] [PubMed]
51. Huerta-Cepas, J.; Serra, F.; Bork, P. ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **2016**, *33*, 1635–1638. [CrossRef] [PubMed]


AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

ENVIRONMENTAL MICROBIOLOGY

# Evolution of *Pectobacterium* Bacteriophage ΦM1 To Escape Two Bifunctional Type III Toxin-Antitoxin and Abortive Infection Systems through Mutations in a Single Viral Gene

**Tim R. Blower,[a]* Ray Chai,[a] Rita Przybilski,[b] Shahzad Chindhy,[a] Xinzhe Fang,[a] Samuel E. Kidman,[a] Hui Tan,[a] Ben F. Luisi,[a] Peter C. Fineran,[b] George P. C. Salmond[a]**

Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom[a]; Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand[b]

**ABSTRACT** Some bacteria, when infected by their viral parasites (bacteriophages), undergo a suicidal response that also terminates productive viral replication (abortive infection [Abi]). This response can be viewed as an altruistic act protecting the uninfected bacterial clonal population. Abortive infection can occur through the action of type III protein-RNA toxin-antitoxin (TA) systems, such as ToxIN$_{Pa}$ from the phytopathogen *Pectobacterium atrosepticum*. Rare spontaneous mutants evolved in the generalized transducing phage ΦM1, which escaped ToxIN$_{Pa}$-mediated abortive infection in *P. atrosepticum*. ΦM1 is a member of the *Podoviridae* and a member of the "KMV-like" viruses, a subset of the T7 supergroup. Genomic sequencing of ΦM1 escape mutants revealed single-base changes which clustered in a single open reading frame. The "escape" gene product, M1-23, was highly toxic to the host bacterium when overexpressed, but mutations in M1-23 that enabled an escape phenotype caused M1-23 to be less toxic. M1-23 is encoded within the DNA metabolism modular section of the phage genome, and when it was overexpressed, it copurified with the host nucleotide excision repair protein UvrA. While the M1-23 protein interacted with UvrA in coimmunoprecipitation assays, a UvrA mutant strain still aborted ΦM1, suggesting that the interaction is not critical for the type III TA Abi activity. Additionally, ΦM1 escaped a heterologous type III TA system (TenpIN$_{Pl}$) from *Photorhabdus luminescens* (reconstituted in *P. atrosepticum*) through mutations in the same protein, M1-23. The mechanistic action of M1-23 is currently unknown, but further analysis of this protein may provide insights into the mode of activation of both systems.

**IMPORTANCE** Bacteriophages, the viral predators of bacteria, are the most abundant biological entities and are important factors in driving bacterial evolution. In order to survive infection by these viruses, bacteria have evolved numerous antiphage mechanisms. Many of the studies involved in understanding these interactions have led to the discovery of biotechnological and gene-editing tools, most notably restriction enzymes and more recently the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas systems. Abortive infection is another such antiphage mechanism that warrants further investigation. It is unique in that activation of the system leads to the premature death of the infected cells. As bacteria infected with the virus are destined to die, undergoing precocious suicide prevents the release of progeny phage and protects the rest of the bacterial population. This altruistic suicide can be caused by type III toxin-antitoxin systems, and understand-

**Received** 28 November 2016 **Accepted** 26 January 2017

**Accepted manuscript posted online** 3 February 2017

**Citation** Blower TR, Chai R, Przybilski R, Chindhy S, Fang X, Kidman SE, Tan H, Luisi BF, Fineran PC, Salmond GPC. 2017. Evolution of *Pectobacterium* bacteriophage ΦM1 to escape two bifunctional type III toxin-antitoxin and abortive infection systems through mutations in a single viral gene. Appl Environ Microbiol 83:e03229-16. https://doi.org/10.1128/AEM.03229-16.

**Editor** Frank E. Löffler, University of Tennessee and Oak Ridge National Laboratory



Address correspondence to George P. C. Salmond, gpcs2@cam.ac.uk.

* Present address: Tim R. Blower, Department of Biosciences, Durham University, Durham, United Kingdom.

ing the activation mechanisms involved will provide deeper insight into the abortive infection process.

**KEYWORDS** type III toxin-antitoxin, ΦM1, *Pectobacterium atrosepticum*, abortive infection, bacteriophage-bacterium interaction

It is estimated that there are more than $10^{30}$ bacteriophages (phages) on Earth, outnumbering their bacterial hosts 10-fold (1, 2). These large viral numbers generate an estimated $10^{25}$ infections per second, imposing a large evolutionary selection pressure on bacteria (2). In response, bacteria have evolved a plethora of defensive mechanisms to counter these overwhelming phage insults (3). Consequently, phages are continually evolving counterdefenses, and thus both the host and parasite are locked together in a perpetual molecular arms race (4). Bacterial antiphage mechanisms that have been observed include adsorption prevention, restriction-modification systems, superinfection systems, abortive infection (Abi) systems, and the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas systems (3). Studies of these phage-host interactions have been translated into significant molecular technologies and reagents, most notably the use of restriction enzymes in cloning (5) and, more recently, the CRISPR-Cas systems, the use of which is currently revolutionizing eukaryotic molecular biology (6).

One of the more curious antiphage mechanisms is Abi, in which, postinfection, the host bacterium is driven toward precocious cell death. This simultaneously terminates viral replication and prevents a productive phage burst. Thus, the Abi response in infected cells protects the bacterial population from progeny phage infection in a process akin to an altruistic suicide (3). The majority of Abi systems have been studied in *Lactococcus lactis* (7), an important bacterium in the dairy industry (8). Phage contamination in fermentation cultures can cause substantial economic losses. Consequently, considerable research has been conducted to identify and define many antiphage systems useful for control of bacteriophages in lactococcal fermentations (7). However, there are also well-studied Abi systems in other bacteria, such as *Escherichia coli*, namely, the Rex, Lit, and PrrC systems (9–11). A commonly recurring theme of Abi systems is that they involve the activation of a toxic protein that is suppressed under normal growth conditions. However, environmental insults, phages, or other physiological stresses can activate the toxin. Once activated, the toxin interferes with an essential cellular process and induces bacteriostasis, ultimately leading to cell death. This is a common feature shared by toxin-antitoxin (TA) systems (12).

TA systems were originally discovered on plasmids, where they function as plasmid maintenance systems through postsegregational killing mechanisms (13). They have been found in the majority of bacteria, both on plasmids (13) and chromosomally (14), as well as in archaea (15) and phages (16). TA systems are typically bicistronic, comprising a bacteriostatic or bactericidal toxic protein that is neutralized either directly or indirectly by an antitoxin counterpart. To date, there are six TA system types which are characterized by the nature and mode of action of their antitoxins (17). In the case of type III TA systems, an RNA antitoxin directly interacts with the toxic protein to form a nontoxic complex (18).

At least four types of TA systems confer phage resistance. These are the *hok/sok* systems of type I (19), *mazEF*, *rnlAB*, and *lsoAB* of type II (20, 21), $ToxIN_{Pa}$, $TenpIN_{Pl}$, and AbiQ of type III (22–24), AbiE of type IV (25), and sanaTA (which is currently not characterized but likely to be a type II, having a proteinaceous antitoxin) (26). $ToxIN_{Pa}$ was the first type III system to be identified and originated from *Pectobacterium atrosepticum* plasmid pECA1039. The toxin $ToxN_{Pa}$ is encoded by *toxN*, and the antitoxin $ToxI_{Pa}$ is encoded by *toxI*, a 36-nucleotide sequence repeated five and a half times (22). The $ToxIN_{Pa}$ system provides protection against multiple phages infecting not only its cognate host, *P. atrosepticum*, but also other enteric bacteria, including *E. coli* DH5$\alpha$ and *Serratia marcescens* Db11 (22). One such aborted pectobacterial phage is the *Myoviridae* phage ΦTE. ΦTE phages that were no longer sensitive to $ToxIN_{Pa}$ had

evolved to encode an RNA antitoxic mimic of $ToxI_{Pa}$, which was able to neutralize $ToxN_{Pa}$ (27). However, it did not shed light on how $ToxIN_{Pa}$ was activated during phage infection. In fact, very little is known about the activation of any type III toxin-antitoxin systems. The other type III system that has been studied for Abi is AbiQ from *Lactococcus lactis*, which shows structural homology with $ToxN_{Pa}$ (24). Three lactococcal siphophages that were aborted by AbiQ have been examined in detail. However, all had mutations in genes of unknown functions; *orf38*, *m1*, and *e19* of phages P008, bIL170, and c2, respectively (28). The AbiQ system was also reconstructed in a heterologous host, *E. coli* MG1655, and was shown to confer resistance to a range of coliphages, including T4 and T5. However, escape mutants could be obtained only for a single phage (phage 2). Escape mutants of this phage showed mutations in *orf210*, a predicted DNA polymerase (28). Studies of the AbiQ system suggests that there may be multiple potential routes of escape involving several genes from different phages in the activation of a single Abi system.

Previously it was shown that the pectobacterial phage ΦM1 was aborted by the $ToxIN_{Pa}$ system and was able to escape by evolving rare mutants (29). ΦM1 was isolated in 1995 during a search for new transducing phages effective as genetic tools in *P. atrosepticum* (30). Here we characterize ΦM1 and its escape mutants in depth. All ΦM1 escape phages evolved through mutations in a gene encoding a small, highly toxic protein, M1-23. When the related $TenpIN_{Pl}$ system of *Photorhabdus luminescens* was transferred to *P. atrosepticum*, the system was able to abort ΦM1 in the heterologous host. Furthermore, it was possible to select spontaneous viral mutants that escaped both $ToxIN_{Pa}$ and $TenpIN_{Pl}$ through mutations in M1-23.

## RESULTS

**ΦM1 is a "KMV-like" virus.** ΦM1 is a generalized transducing phage of *Pectobacterium atrosepticum* (previously *Erwinia carotovora* subsp. *atroseptica*) (30). This podovirus is aborted by the type III TA system, ToxIN, from *P. atrosepticum*, namely, $ToxIN_{Pa}$ (22). ΦM1 generates spontaneous escape mutants that are resistant to Abi by $ToxIN_{Pa}$ at a rate of $\sim 10^{-5}$ (29). In order to improve our understanding of $ToxIN_{Pa}$-phage interactions, we sequenced ΦM1 wild type (wt) and three previously isolated escape phages, ΦM1-A, -B, and -D (29).

Using BLAST searches (31), ΦM1 was classified as a member of the "KMV-like" subgroup of the T7 supergroup of phages (32). T7-like phage linear genomes are typically flanked by direct terminal repeats (DTRs) (33). However, the DTRs could not be defined by a primer walking strategy along the ΦM1 genome, consistent with results from another KMV-like phage, LIMEzero (34). The presence and approximate size of the DTRs, 293 bp, were therefore confirmed through restriction digest analysis of the ΦM1 genome (see Fig. S2 in the supplemental material). The final ΦM1 wild-type genome was 43,827 bp long with a GC content of 49.30%. In comparison, the host *P. atrosepticum* genome has a GC content of 50.97% (35). The two genomes therefore closely match each other in GC content.

Global nucleotide alignments were performed to assess the relationship between the KMV-like phages and ΦM1. Compared with ΦM1, phage VP93 (43,931 bp) (36), phage LKA1 (41,593 bp) (32), phage LKD16 (43,200 bp) (32), and ΦKMV itself (42,519 bp) (33) shared between 48.2% and 49.2% sequence identity. These values match well those of other KMV-like phages (34).

ΦM1 contains 52 putative genes, named *phiM1-1* to *phiM1-52*. The gene products were named M1-1 to M1-52, and they are encoded by 92.6% of the genome. Subsequent BLASTp searches identified homologues for 32 of the open reading frames (ORFs) from other KMV-like phages (Table S1). In most cases, it was therefore possible to assign putative functions and categorize ORFs as containing either metabolism, structural, or host lysis genes (Fig. 1A). ΦM1 also encodes a single $tRNA^{Ile}$, between *phiM1-38* and *phiM1-39*.

**ΦM1 escape mutations had specific base substitutions.** The genome sequences of the three escape phages, M1-A, -B, and -D, were compared with that of the wt. All

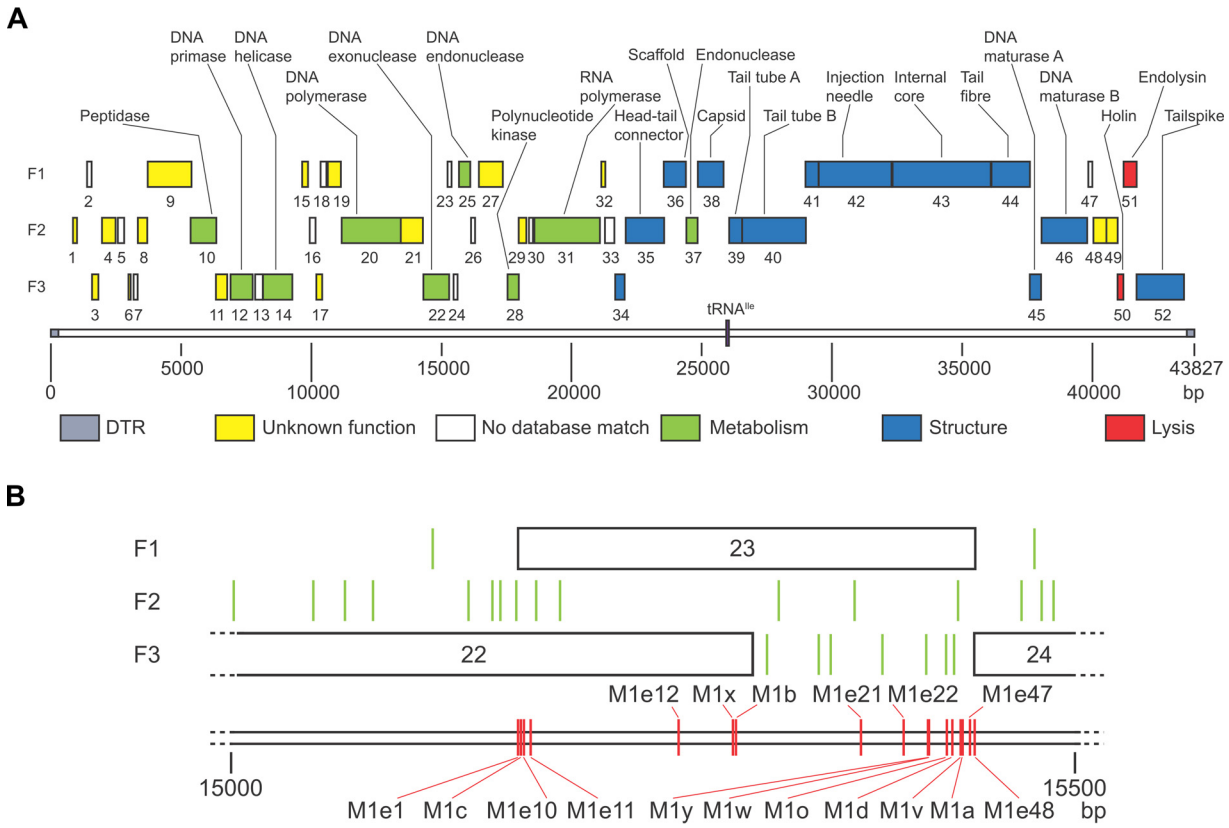



**FIG 1** Genomic map of ΦM1 wild type and its escape locus. (A) All 52 annotated ORFs are coded on the forward reading strand, in a linear progression from metabolic genes to structural genes and, finally, to host cell lysis genes. Each forward reading frame is labeled F1, F2, or F3. ORFs are shown to scale as shaded boxes numbered with the gene number, colored according to the predicted role. The single $tRNA^{Ile}$ gene is positioned on the scale, shown in purple. Where it was possible to identify a protein by homology searches, that ORF is labeled. The scale is in base pairs. The figure was drawn to scale using Adobe Illustrator. (B) Schematic of the escape locus of ΦM1. All escape phage mutations are within *phiM1-23*. Each forward reading frame is labeled F1, F2, or F3. Each ORF is shown to scale as a box, numbered with the gene number. Each stop codon is represented as a green vertical line. The positions of the ΦM1 escape phage mutations are shown by red vertical lines, labeled with the parent phage. The scale is in base pairs.

three escape phages had single point mutations localized to a 124-bp stretch (Fig. 1B), across *phiM1-22* and *phiM1-23*, which we refer to as the "escape locus." To ascertain whether these point substitutions were individual changes, further escape phages were isolated using independent lysates to avoid the possibility of sibling mutants. The new escape phage mutants were isolated following selection on *P. atrosepticum* pTA46 ($ToxIN_{Pa}$) (22, 29). The escape locus of each phage was sequenced following amplification of the region from the purified genomic DNA. We observed that all 10 escape phages had unique mutations distributed across 246 bp of the escape locus (Fig. 1B). Nine of these mutations were base substitutions, while one was a single base deletion (Table 1).

**Infection with ΦM1 affects the $ToxI_{Pa}$/$ToxN_{Pa}$ ratio.** Though it has been shown that $ToxN_{Pa}$ levels do not alter during a ΦM1 phage infection (29), it was not known how the $ToxI_{Pa}$ levels were affected. The identification of the escape phages provided an opportunity to address this question. To investigate alterations to the $ToxI_{Pa}$/$ToxN_{Pa}$ ratio, we monitored the levels of $ToxI_{Pa}$ and $ToxN_{Pa}$-FLAG during the infections by ΦM1 and the escape phage ΦM1-O within *P. atrosepticum* carrying a $ToxIN_{Pa}$-FLAG plasmid (pMJ4). Total protein and RNA samples were taken at different times after infection and subjected to Western blotting and an S1 nuclease assay, respectively. While $ToxN_{Pa}$ levels stayed constant throughout infection (Fig. 2A, lower panel), $ToxI_{Pa}$ levels dropped dramatically after 30 min compared to those of an uninfected control (Fig. 2A). Interestingly, $ToxI_{Pa}$ levels increased back to original levels at 60 min. In comparison to the infection with ΦM1 wt, $ToxI_{Pa}$ levels did not change significantly at 30 min during

**TABLE 1** Summary of ΦM1 escape mutations and effects on reading frames

| Phage | Date of isolation | Position and mutation relative to ΦM1 wt[a] | Effect on forward reading frame[b]: | | |
|---|---|---|---|---|---|
| | | | F1 | F2 | F3 |
| ΦM1-A | March 2007 | 15416, A to C | Y to S | T to P | No change |
| ΦM1-B | March 2007 | 15292, C to T | R to stop | No change | P to L |
| ΦM1-C | March 2007 | 15170, T to C | M to T | Stop to S | No change |
| ΦM1-D | March 2007 | 15410, T to C | M to T | W to R | No change |
| ΦM1-O | June 2009 | 15407, A to C | Q to P | No change | No change |
| ΦM1-V | May 2009 | 15415, T to G | Y to D | No change | V to G |
| ΦM1-W | May 2009 | 15398, A to T | D to V | M to L | Stop to C |
| ΦM1-X | May 2009 | 15288, AA to A | FS to stop after 9 aa (wild-type F1 continues *phiM1-23*) | FS causing Q to H and stop after 3 aa (wild-type F2 stops after 9 aa) | FS causing N to T and shift of ORF1 into ORF2 (wild-type F3 stops after 3 aa) |
| ΦM1-Y | May 2009 | 15397, G to A | | | |
| ΦM1-Z | May 2009 | 15416, A to G (cf. ΦM1-A) | | | |

[a]Mutations are indicated as, e.g., Y to S.
[b]FS, frameshift; aa, amino acid.

infection with the escape phage ΦM1-O (Fig. 2B). The $ToxI_{Pa}$ level did decrease with the ΦM1-O infection but only at 40 min (Fig. 2B). The $ToxI_{Pa}$ levels were not then restored, as in the case of ΦM1 wt (Fig. 2B). ΦM1 appears to activate $ToxN_{Pa}$, and thereby initiate Abi, by causing a decrease in the cellular $ToxI_{Pa}$ levels, either through direct or indirect means. In the case of ΦM1-O, this activation is prevented due to the mutation in M1-23. This would allow the phage to propagate, which may then account for the delayed decrease and lack of restoration in $ToxI_{Pa}$ levels.

**Identification and characterization of the ΦM1 escape product.** The majority of escape mutations occurred within *phiM1-23*. On first analysis, two mutations, those from ΦM1-B and ΦM1-X, occurred at the 3′ end of *phiM1-22*. Another mutation, from ΦM1-C, mapped further upstream, again within *phiM1-22*. This gene, *phiM1-22*, encodes a homologue of a putative DNA exonuclease from phage LKA1 (Table S1) (32). Unfortunately, there were no database hits for *phiM1-23* and *phiM1-24*, using either the nucleotide or encoded protein sequences.

Specific regions of this escape locus were amplified from ΦM1 phages and then cloned into pBAD30 (37) to make inducible constructs (Fig. 3A and B). The cloning began with constructs 1 to 6, using DNA from ΦM1 wt and ΦM1-B (Fig. 3B). Constructs 1 and 2 could not be obtained with ΦM1 wt DNA, presumably through toxicity of the resulting wt constructs in *E. coli* DH5α, but could be made using

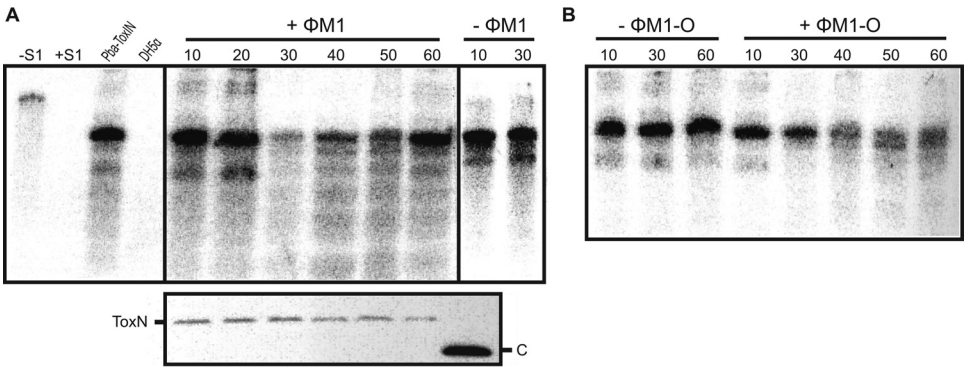



**FIG 2** $ToxI_{Pa}$ levels are affected during phage infection. (A) S1 nuclease assay targeting the full 5.5-repeat $ToxI_{Pa}$ sequence was used to monitor $ToxI_{Pa}$ levels during ΦM1 infection. Assays were performed on 10 μg total RNA prepared from *P. atrosepticum* $ToxIN_{Pa}$ (pMJ4) at different times following ΦM1 infection. Numbers indicate the time (minutes) after infection with phage (+ΦM1) and the negative control without phage (−ΦM1). Hybridization to total RNA from *P. atrosepticum* expressing $ToxIN_{Pa}$ (pTA46) and DH5α served as positive and negative controls, respectively. The expression of $ToxN_{Pa}$ at the respective time points of infection is shown in the lower panel using Western blotting; "C" indicates the 11-kDa SdhE-FLAG protein used as a loading and size control (54). (B) S1 nuclease assay targeting $ToxI_{Pa}$ for the infection with the escape phage ΦM1-O. The assay was done as described for panel A.

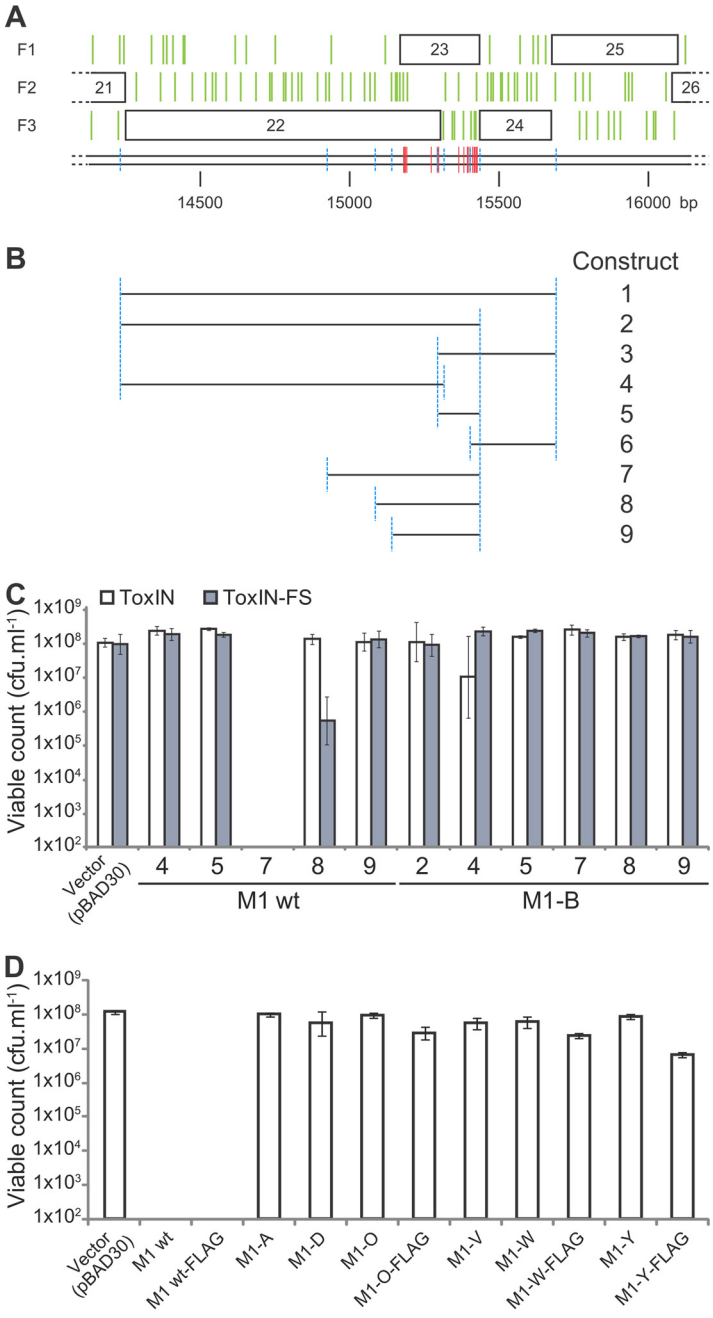



**FIG 3** Toxicity of the ΦM1 escape locus products. (A) The escape locus of ΦM1 as described in the legend to Fig. 1B. The positions of the ΦM1 escape phage mutations are shown by red vertical lines, labeled with the parent phage. The scale is in base pairs. (B) Specific regions of the phage genomes, designated by the length of the line that corresponds to the genomic locus shown in panel A, were cloned into pBAD30 to make nine different constructs. Blue dashed lines in panel A reflect the construct boundaries in panel B. The figure is drawn to scale. (C) Expression of ΦM1 wt and ΦM1-B escape loci in *P. atrosepticum*. Strains of *P. atrosepticum* containing either a $ToxIN_{Pa}$ or $ToxIN_{Pa}$-FS plasmid (pTRB125 or pTRB126), together with a phage construct (or pBAD30 vector control), were tested for toxicity. (D) A range of construct 7 plasmids was tested for toxicity in *P. atrosepticum*. The escape phage constructs were all reduced for toxicity. Error bars show the standard deviations for triplicate data.

ΦM1-B escape phage DNA. Constructs 3, 4, 5, and 6 could be made using both sources of DNA. Due to the regions covered by these constructs, we could determine that within this locus, the genes of interest were *phiM1-22* and *phiM1-23* and that *phiM1-24* did not contribute to toxicity. As pBAD30 is tightly repressed by

glucose in *E. coli* DH5α, this also implied that toxicity from this region of DNA might be occurring via an internal promoter.

Upon first analysis, the putative ATG start of *phiM1-23* was at bp 15304. Taking into account the stop codons of each frame (Fig. 1B, green vertical lines), the putative ATG start codon of *phiM1-23* could theoretically have been upstream of this initial annotation. There were three possible ATG sites upstream of the putative start codon for *phiM1-23*. The mutation of ΦM1-C specifically altered the middle of these start codons from M to T (Table 1). This start codon also had a ribosome binding site closer to consensus than those of the other potential start codons, making it the most obvious candidate. If this were the case, the escape mutations would span *phiM1-23* specifically. Constructs 7 to 9 were designed and made in order to test whether *phiM1-23* alone could generate a toxic phenotype.

We performed experiments to assess the toxicity of the escape locus constructs and to determine whether toxicity was related to the presence of ToxIN$_{Pa}$. *P. atrosepticum* was transformed with inducible derivatives of the escape locus in combination with either pBluescript-based (Fermentas) ToxIN$_{Pa}$ or negative-control ToxIN$_{Pa}$-frameshift (FS) vectors (pTRB125 and pTRB126, respectively). Serial dilutions of these dual-vector strains of *P. atrosepticum* were then incubated with and without induction, overnight, to determine the viable count (Fig. 3C). This clearly showed that the product of construct 7, covering *phiM1-23* specifically, was toxic. There was no toxicity in the case of ΦM1-B, the mutation in which causes a premature stop codon in *phiM1-23*. Toxicity was also independent of the presence of ToxIN$_{Pa}$. These results strongly suggested that *phiM1-23* produces a small, toxic protein, responsible either directly or indirectly for activation of Abi against ΦM1.

New versions of construct 7 (Fig. 3D) were then generated, with the addition of a C-terminal FLAG tag to the M1-23 product, using both ΦM1 wt and escape sequences. Various constructs were then tested for toxicity in the cognate host, *P. atrosepticum* (Fig. 3D). All the escape constructs tested showed reduced toxicity (Fig. 3D). It was therefore possible to attempt overexpression and purification of M1-23, using an *E. coli* expression strain, ER2566. After expression trials using constructs made from ΦM1 wt and ΦM1-O, -W, and -Y phage DNA, the M1-O-23FLAG product was chosen for further study. Sufficient M1-O-23FLAG protein was purified to allow mass spectrometry to confirm both the identity of the protein and, specifically, the presence of the expected Q-to-P mutation. Furthermore, the protein sample was subjected to N-terminal sequencing, generating a sequence of TKM. This implied that *phiM1-23* started at the ATG specifically mutated by ΦM1-C, as described earlier, and that the initial methionine is cleaved posttranslationally. The annotation of the ΦM1 wt genome was then altered to accommodate *phiM1-23* beginning at this confirmed start codon. In summary, this result shows that all the escape mutations map to a single gene, *phiM1-23*, which generates a 9.8-kDa protein. These mutations reduce the toxicity of the protein product and allow viral escape from ToxIN$_{Pa}$-induced Abi.

It had not been possible to clone constructs 1 and 2 (Fig. 3B) using the ΦM1 wt sequence, despite the pBAD30 vector system being repressed in the presence of glucose. This suggested that a promoter internal to those cloned regions might be inducing the transcription of *phiM1-23*. A range of pRW50-based (38) *lacZ* transcriptional fusion constructs was generated to investigate the possible presence of a promoter (Fig. S3A). In this case, it was possible to clone the equivalent of construct 2 using ΦM1 wt DNA (Fig. 3B), perhaps due to pRW50 having a low copy number, so the level of toxicity was sufficiently low. Plasmid pTA104 (22), containing the promoter for ToxIN$_{Pa}$, was used as a positive control. All the test constructs except pTRB162, which was an extremely truncated clone, generated LacZ activity (Fig. S3B). This confirmed the presence of a weak *phiM1-23* promoter within *phiM1-22*.

**Extensive analysis of ΦM1 escape mutants map all mutations to *phiM1-23*.** The initial 10 escape mutants of ΦM1 all had unique mutations in M1-23, so it was likely that there were other possible mutations not yet observed. Identifying these other mutations

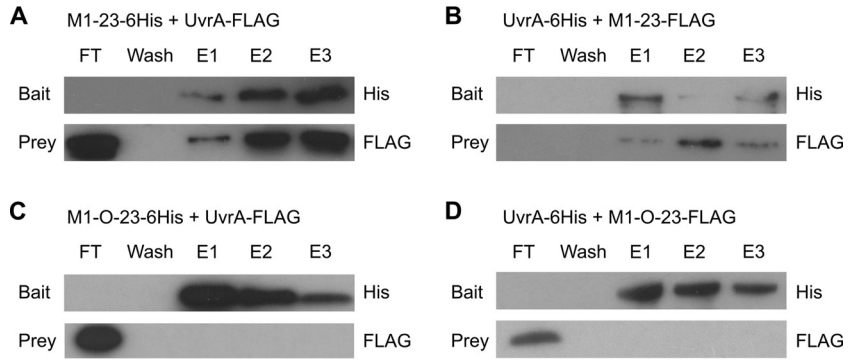



**FIG 4** Coimmunoprecipitation of M1-23, M1-O-23, and UvrA. (A and B) Coimmunoprecipitation experiments with wild-type M1-23 and UvrA. (A) M1-23-6His was used as the bait and attached to a $Ni^{+}$ column with UvrA-FLAG passed through. (B) The reciprocal experiment was performed with UvrA-6His used as the bait with M1-23-FLAG passed through. (C and D) The same coimmunoprecipitation experiments as described for panels A and B but using M1-O-23 instead of M1-23. (C) M1-O-23 was used as bait; (D) UvrA-6His was used as bait.

could reveal important residues involved in the functionality of M1-23. Consequently, a larger library of escape mutants was isolated and characterized in the same way as the initial escape mutants. A total of 51 new, independent escape phages were isolated, and their *phiM1-23* sequences were characterized. All escape phages were shown to have a mutation in this region, and several new unique escape phages were isolated (Table S2). With the addition of these new escape phages, the number of different mutations increased to 20. Interestingly, mutations in all three of the bases of the putative start codon were isolated, consistent with this being the correctly annotated start site. Other interesting mutations were those causing N-terminally located truncations of M1-23. In particular, ΦM1-E11 produced only a hypothetical dipeptide or indeed just a single amino acid if the initial starting methionine was removed. Although most mutations in M1-23 were missense alleles generating single amino acid residue changes, the ability to isolate derivatives with major truncations showed that the M1-23 protein must be nonessential for a productive ΦM1 lytic cycle. Other notable mutations were ΦM1-E48 and ΦM1-E49 (both generating the same outcome), which modify the stop codon and lead to a 10-amino-acid C-terminal extension. It is puzzling why the 10-mer extension might impact function, because the addition of the octameric FLAG tag to the C terminus of M1-23 did not disrupt protein toxicity. Perhaps the extension might harbor a sequence that could act as an autoinhibitor or disrupt protein structure.

**M1-23 interacts with UvrA, but abortive infection can still take place in UvrA-deficient *P. atrosepticum*.** To assess whether there is a direct interaction of M1-23 with the $ToxIN_{Pa}$ complex, His-tagged forms of both M1-23 and M1-O-23 were cloned, allowing overexpression and purification of these proteins. Coimmunoprecipitation reactions were carried out, but the results showed no evidence for interactions between M1-23 and the $ToxIN_{Pa}$ complex and no impact of M1-23 on the ToxI RNA (data not shown).

During the process of purifying M1-23-6His, it was noted that an additional high-molecular-weight band that was not present in control samples appeared in the eluted sample, and it was then copurified with M1-23 following ion-exchange fast performance liquid chromatography (FPLC; data not shown). Mass spectrometric analysis identified the host nucleotide excision repair protein, UvrA. Reciprocal coimmunoprecipitation assays were performed using purified protein samples to confirm this interaction (Fig. 4). M1-23 protein retained UvrA, while M1-O-23 did not, and similarly, only M1-23 was retained by immobilized UvrA (Fig. 4). This strongly suggests that M1-23 is a viral product that is able to bind host UvrA.

To assess potential effects of UvrA on abortive infection, a *uvrA* mutant was constructed in *P. atrosepticum* and confirmed by sequencing and then by hypersensi-

**TABLE 2** EOPs against $ToxIN_{Pa}$ and $TenpIN_{Pl}$ type III TA systems

| Phage | EOP vs $ToxIN_{Pa}$ | EOP vs $TenpIN_{Pl}$ | System on which escape was selected |
|---|---|---|---|
| ΦM1 wt | $1.3 \times 10^{-5}$ | $1.1 \times 10^{-2}$ | |
| ΦS61 | $<3.2 \times 10^{-9}$ | 0.9 | |
| ΦTE | $1.0 \times 10^{-8}$ | 0.7 | |
| ΦM1-O | 1.0 | 1.0 | $ToxIN_{Pa}$ |
| ΦM1-PL2 | 0.9 | 0.9 | $TenpIN_{Pl}$ |

tivity to UV light (Fig. S1). This strain was tested for its ability to abort ΦM1 via the $ToxIN_{Pa}$ system. Surprisingly, ΦM1 was still aborted in the *uvrA* mutant and to the same extent as in the wild-type *P. atrosepticum* strain (efficiency of plating [EOP] of ΦM1 on the *uvrA* mutant with $ToxIN_{Pa}$, $1.1 \times 10^{-5}$). Escape phages of ΦM1 were isolated from the *uvrA* mutant, and their DNA was sequenced. Interestingly, all escape phages isolated on the *uvrA* mutant, ΦM1-U1, -U2, and ΦM1-U4 to ΦM1-U10 (which were independently isolated), carried mutations in the M1-23 sequence (Table S2). The results suggest that although M1-23 clearly has a specific interaction with UvrA, it appears that the escape route is either subtle or occurs indirectly.

**The ΦM1 escape mechanism works in another type III TA and Abi system.** Two further families of type III TA systems were recently identified, CptIN and TenpIN (23). $TenpIN_{Pl}$, from the chromosome of *Photorhabdus luminescens* TT01, was able to act as an Abi system against coliphages when cloned on a multicopy plasmid and tested in an *E. coli* background (23). By transforming *P. atrosepticum* SCRI1043 with the $TenpIN_{Pl}$ expression plasmid, pFR2 (23), we were able to test three *Pectobacterium* phages against the Abi activity of $TenpIN_{Pl}$ (Table 2). While ΦS61 (29) and ΦTE (27) were dramatically affected by $ToxIN_{Pa}$, neither were inhibited by $TenpIN_{Pl}$ (Table 2). This indicates a degree of selectivity between the two Abi systems. ΦM1, however, was aborted by both systems, though to different degrees, which also underlines the selectivity under which $ToxIN_{Pa}$ and $TenpIN_{Pl}$ appear to operate. As with $ToxIN_{Pa}$, it was possible to select for phages of ΦM1 that escaped Abi by $TenpIN_{Pl}$. One of these escape phages, ΦM1-PL2, was isolated and sequenced. This escape phage had a single base substitution, T15410C, the same mutation as in ΦM1-D. To test this in reverse, escape phage ΦM1-O, selected with $ToxIN_{Pa}$, was tested against $TenpIN_{Pl}$ (Table 2). ΦM1-O was also resistant to $TenpIN_{Pl}$. These results imply that in the case of ΦM1, the two systems operate in a similar fashion with a single protein, M1-23, being a key mediator.

## DISCUSSION

The pectobacterial phage ΦM1 was shown previously to be sensitive to the $ToxIN_{Pa}$ system and capable of producing spontaneous escape mutants (29). Here we found that the ΦM1 phage is also sensitive to $TenpIN_{Pl}$ when reconstructed in *P. atrosepticum* and is correspondingly able to evolve escape mutants. This is the first time we have been able to identify a phage that is able to escape the $TenpIN_{Pl}$ system, and so further study may provide information about its activation. Interestingly, the ΦM1 phage is insensitive to two other type III systems tested, $ToxIN_{Bt}$ from *Bacillus thuringiensis* and the $CptIN_{Er}$ system from *Eubacterium rectale* (data not shown), and no Abi activity has so far been observed in these two systems (23, 39). In contrast, the *P. atrosepticum* phage ΦTE is aborted by $ToxIN_{Pa}$ and able to escape the system by RNA-based molecular mimicry of the antitoxin (27) but is not aborted by the $TenpIN_{Pl}$ system (Table 2).

Characterization of the ΦM1 phage in this study has shown that all escape mutants selected on $ToxIN_{Pa}$ or $TenpIN_{Pl}$ have mutations in a gene encoding M1-23. Alteration of single amino acids, extreme truncations due to very 5′ stop codons, or even stop codon mutations leading to short C-terminal extensions of M1-23 cause insensitivity to both $ToxIN_{Pa}$ and $TenpIN_{Pl}$. Escape mutants selected on one system are also insensitive to the other system, suggesting that there is a common pathway for the ΦM1 phage in the activation of these two systems. The role of M1-23 is unknown, but it was shown

to be nonessential, and as it is located between a predicted exonuclease gene, *phiM1-22*, and a predicted endonuclease gene, *phiM1-25*, it might have a role in the regulation of nucleases or indeed may be able to act as a nuclease itself. In a previous study, it was shown that $ToxN_{Pa}$ levels do not change during infection of the ΦM1 phage (29). In this study, we found that the $ToxI_{Pa}$ levels decrease 30 min postinfection. In contrast, during the infection by the ΦM1 escape phage ΦM1-O, $ToxI_{Pa}$ levels decreased only slightly after 40 min and were not restored. It appears that wild-type ΦM1 activates $ToxN_{Pa}$ by decreasing the levels of $ToxI_{Pa}$ and therefore initiating Abi. For ΦM1-O, the mutation in M1-23 prevents this early activation and thereby provides a window of opportunity for the phage to replicate.

To investigate the mechanism of M1-23 action, a large number of ΦM1 escape phages were isolated and their *phiM1-23* regions were sequenced. The results showed a number of escape mutations near the 5′ end of the gene, resulting in extremely truncated versions of the protein. This confirms that M1-23 is a nonessential viral protein. However, the majority of mutations found were toward the 3′ region of the gene and were mostly missense mutations resulting in single amino acid changes, implying that the C-terminal domain of the protein is important for Abi functionality. To further characterize M1-23, it was overexpressed and purified, but due to high toxicity, only a small amount of protein could be produced. Using the limited amount of protein available, interaction studies were performed to see if M1-23 interacted with $ToxIN_{Pa}$. During purification of M1-23, a high-molecular-weight protein always copurified. Mass spectrometry of this protein confirmed that it was the DNA repair protein UvrA. It was shown through coimmunoprecipitation experiments that while M1-23 could interact with UvrA, the escape version of the protein M1-O-23 could not.

UvrA forms part of the SOS response in bacteria, a DNA damage response pathway (40) that has previously been shown to be involved in TA activation. The type I TA system TisB-IstR is under direct SOS response control, as *tisAB*, which encodes the TisB toxin, contains a LexA operator region that is inhibited by LexA (41). In addition to the SOS response, the stringent response has also been shown to play a role in the activation of TA systems. Both type I and type II TA systems have been shown to be regulated by (p)ppGpp, the central regulator of the stringent response (42, 43). However, ΦM1 and ΦTE were tested in a (p)ppGpp-negative double mutant (*relA spoT*) and were still aborted in that background (data not shown).

During the course of this study, the genomes of two new pectobacterial phages were sequenced. These were *P. atrosepticum* phage Peat1 (44) (GenBank accession number KR604693) and *P. carotovorum* phage PPWS1 (45) (DDBJ accession number LC063634). Both of these were podoviruses that shared high sequence identity to ΦM1. Peat1 (45,633 bp) shared 77.7% sequence identity, and PPWS1 (44,539 bp) shared 59.7% sequence identity. Furthermore, analysis of the two genomes revealed that both phages encoded M1-23 homologs, with the Peat1 homolog differing by only a single amino acid. Therefore, it is highly likely that both phages would be aborted by both the $ToxIN_{Pa}$ and $TenpIN_{PI}$ systems and evolve escapes in the same way. If this was the case, it would show a common route through which phages of different bacteria are able to escape the same system.

Both $ToxIN_{Pa}$ and $TenpIN_{PI}$ are very powerful antiphage abortive infection systems that belong to two different families of type III TA systems and are effective against a wide variety of phages. While many phages show differing sensitivities to the two systems, this study has shown that in ΦM1, there is a common pathway through which these two families of type III TA systems can be activated. This pathway involves a small toxic protein, M1-23, of unknown metabolic function that does not directly interact with the $ToxIN_{Pa}$ complex but that interacts directly with UvrA. ΦM1 infection causes a diminution in $ToxI_{Pa}$ levels, presumably leading to the destabilization of the $ToxIN_{Pa}$ complex and consequent liberation of $ToxN_{Pa}$ to induce cell death and concomitant abortive infection of the viral parasite.

**TABLE 3** Bacterial strains and bacteriophages used in this study

| Bacterium or phage | Genotype or characteristics | Reference or source |
|---|---|---|
| Bacteria | | |
| *Escherichia coli* β2163 | F⁻ RP4-2-Tc::Mu *dapA*::(*erm-pir*) Km$^r$ | 57 |
| *E. coli* DH5α | F⁻ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 hsdR17*($r_K^-$ $m_K^+$) λ⁻ | Gibco/BRL |
| *E. coli* ER2566 | F⁻ λ⁻ *fhuA2* [*lon*] *ompT lacZ*::T7 gene 1 *gal sulA11* Δ(*mcrC-mrr*)114::IS*10* R(*mcr-73*::miniTn*10*-TetS)2 R(*zgb-210*::Tn*10*) (TetS) *endA1* [*dcm*] | NEB |
| *E. coli* W3100 | F⁻ λ⁻ *rph-1* INV(*rrhD*, *rrhE*) | 58 |
| *Pectobacterium atrosepticum* SCRI1043 | Wild-type strain | 35 |
| Phages | | |
| ΦM1 | *Podoviridae*, propagated on wt SCRI1043 | 30 |
| ΦM1-A | ToxIN$_{Pa}$ escape mutant of ΦM1 | 29 |
| ΦM1-B | ToxIN$_{Pa}$ escape mutant of ΦM1 | 29 |
| ΦM1-C | ToxIN$_{Pa}$ escape mutant of ΦM1 | 29 |
| ΦM1-D | ToxIN$_{Pa}$ escape mutant of ΦM1 | 29 |
| ΦM1-O | ToxIN$_{Pa}$ escape mutant of ΦM1 | This study |
| ΦM1-V | ToxIN$_{Pa}$ escape mutant of ΦM1 | This study |
| ΦM1-W | ToxIN$_{Pa}$ escape mutant of ΦM1 | This study |
| ΦM1-X | ToxIN$_{Pa}$ escape mutant of ΦM1 | This study |
| ΦM1-Y | ToxIN$_{Pa}$ escape mutant of ΦM1 | This study |
| ΦM1-Z | ToxIN$_{Pa}$ escape mutant of ΦM1 | This study |
| ΦM1-Q | ToxIN$_{Pa}$ escape mutant of ΦM1 | This study |
| ΦM1-E1 to -E49 | ToxIN$_{Pa}$ escape mutant of ΦM1 | This study |
| ΦM1-U1 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-U2 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-U4 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-U5 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-U6 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-U7 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-U8 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-U9 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-U10 | ToxIN$_{Pa}$ escape mutant of ΦM1 on UvrA mutant | This study |
| ΦM1-PL2 | TenpIN$_{Pl}$ escape mutant of ΦM1 | This study |

## MATERIALS AND METHODS

**Bacterial strains, bacteriophages, and growth conditions.** Bacterial strains and bacteriophages are listed in Table 3. *E. coli* strains were grown at 37°C, and *Pectobacterium atrosepticum* SCRI1043 (35) was grown either at 25°C on agar plates or at 25, 28, or 30°C as required for liquid culture in Luria broth (LB) at 250 rpm or on LB agar (LBA). LBA contained 1.5% (wt/vol) or 0.35% (wt/vol) agar, to make LBA plates or top-LBA, respectively. Bacterial growth was measured using a spectrophotometer set to 600 nm. When required, media were supplemented with ampicillin (Ap) at 100 μg ml$^{-1}$, chloramphenicol (Cm) at 50 μg ml$^{-1}$, kanamycin (Km) at 50 μg ml$^{-1}$, tetracycline (Tc) at 10 μg ml$^{-1}$, isopropyl β-D-thiogalactopyranoside (IPTG) at 0.5 mM, or 2, 6-diaminopimelic acid (DAPA) at 300 μM. Spontaneous escape phage mutants were isolated as described previously (27). Phage lysates were made as described previously (46). Phages were stored at 4°C in phage buffer, i.e., 10 mM Tris-HCl (pH 7.4), 10 mM $MgSO_4$, and 0.01% (wt/vol) gelatin. A few drops of chloroform saturated with sodium bicarbonate was also added to the phage lysates to maintain sterility. EOP was calculated after overnight incubation of serial dilutions of phage lysates in a top-LBA lawn of each bacterial host and recorded as the number of PFU on the test strain relative to the number of PFU on the control strain. EOPs were calculated using *P. atrosepticum* wt or a frame-shifted *toxN* plasmid strain as the negative control (22).

**ΦM1 genomic sequencing.** Bacteriophage DNA was extracted with phenol-chloroform, using Phase Lock Gel tubes (Eppendorf) and in accordance with the manufacturer's instructions, as for bacteriophage λ. The extracted DNA was subjected to pyrosequencing on a Roche 454 Genome Sequencer FLX at the DNA sequencing facility, Department of Biochemistry, University of Cambridge. Contiguous read segments (contigs) were assembled using Newbler (Roche). The ΦM1 wild-type sequence was determined in one lane of the sequencing run. The three escape phage genomes were individually tagged with independent identifying sequences and then combined and sequenced as a mixture within a second lane. For each of the four phages, the final assembled sequence consisted of a single contig of approximately 43,500 bp. The average read length was 250 bp. The wild-type sequence was assembled from 13,628 reads, leading to approximately 78× coverage of the full sequence. Escape phage ΦM1-A, -B, and -D sequences were assembled from 4,925, 5,188, and 5,886 reads, respectively, resulting in approximately 29× coverage of each sequence.

When the sequence data are viewed, beginning at bp 43572 (in the final ΦM1 wt sequence), there are 15 tandem repeats of the 2-bp sequence TG. The number of TG repeats varied between the raw

**TABLE 4** Primers used in this study

| Primer | Sequence (5′–3′) | Description | Restriction site |
|---|---|---|---|
| KDOI | TTTTGGATCCGTTTTATCGACATTGTGAACC | *toxIN* locus | BamHI |
| PF147 | GTATCTAGAGTAGTCGCCTCTTTTACTTTATTAC | *toxI* | XbaI |
| PF217 | TTGTATACTTAAGTTATTGACTCTATAGCTCAG | ToxI amplification for S1 nuclease protection assay | HindIII |
| PF218 | TTGACTATGTAGTCGCCTCTTTTACTTTATTTCGAACCTCGGACCTGCG | ToxI amplification for S1 nuclease protection assay | DrdI |
| TRB37 | CCGGCATATGAAATTCTACACTATATCAAGC | Used for ToxIN CBD | NdeI |
| TRB38 | GTGGTTGCTCTTCCGCACTCGCCTTCTTCCGTAT | Used for ToxIN CBD | SapI |
| TRB107 | TTGAATTCTGCGCAAGCAACTGGTGCACC | ΦM1 sequencing primer | EcoRI |
| TRB108 | TTAAGCTTCTTGAATCTGTACTCACCG | ΦM1 sequencing primer | HindIII |
| TRB111 | TTGAATTCCTGTAGGAGCGTGGAATGC | ΦM1 escape locus | EcoRI |
| TRB115 | TTGAATTCCAGGGGTGTTACCTACTCC | ΦM1 sequencing primer | EcoRI |
| TRB116 | TTAAGCTTGTAACTGTGCAGTGATACC | ΦM1 sequencing primer | HindIII |
| TRB117 | TTGAATTCCCTACAATGCCCCAGATGC | ΦM1 escape locus | EcoRI |
| TRB118 | TTAAGCTTACGGTCGTACTTGGCTTCG | ΦM1 escape locus | HindIII |
| TRB125 | TTAAGCTTCTAATCCTACGCCTTGTGC | ΦM1 escape locus | HindIII |
| TRB126 | TTGAATTCAAGGTGGATGCAACTCGGG | ΦM1 escape locus | EcoRI |
| TRB127 | TTAAGCTTCTCTACATCATCCAACATC | ΦM1 escape locus | HindIII |
| TRB128 | TTGAATTCGAGCTGCGTGATGAGTTCC | ΦM1 escape locus | EcoRI |
| TRB129 | TTGAATTCGCTTACCCGATTATATCC | ΦM1 escape locus | EcoRI |
| TRB130 | TTGAATTCCCAATTTAAAATTAATGA | ΦM1 escape locus | EcoRI |
| TRB134 | TTAAGCTTATTACTTGTCATCGTCGTCCTTGTAGTCTCCTAGGTACCCCATCTGG | ΦM1 construct 7/ORF23 FLAG | HindIII |
| TRB135 | TTAAGCTTAGTGATGGTGATGGTGATGTCCTCCTAGGTACCCCATCTGG | ΦM1 construct 7/ORF23-6His | HindIII |
| TRB332 | TTAAGCTTATTACTTGTCATCGTCGTCCTTGTAGTCTCCCAGCATCGGCTTAAGGAAGCG | *uvrA*-FLAG | HindIII |
| TRB337 | ATTAGGATCCGATAAGATCGAAGTTCG | *uvrA* primer | BamHI |
| TRB338 | ATTAAAGCTTTTACAGCATCGGCTTAAG | *uvrA* primer | HindIII |
| UvrA dnF | TTTATTCCGGGAAGTGTGTGAATTTAAATTAGCGAGAGGCCAAATCATG | Fwd, 500 bp downstream of *uvrA* | SwaI |
| UvrA dnR | TTATCAGAATTCCTGCCGTGCAGGCAGTTCAG | Rev, 500 bp downstream of *uvrA* | EcoRI |
| UvrA upF | TTATCATCTAGATTGCAGTGCGCCTTCGATG | Fwd, 500 bp upstream of *uvrA* | XbaI |
| UvrA upR | CATGATTTGGCCTCTCGCTAATTTAAATTCACACACTTCCCGGAATAAA | Rev, 500 bp upstream of *uvrA* | SwaI |

sequences of each phage, from 17 in ΦM1-A to 1 in ΦM1-B and 7 in ΦM1-D. The exact number of TG repeats in each phage genome could not be accurately confirmed by sequencing a specific amplicon. Therefore, in order to sequence this region, it was specifically amplified (primers TRB107/TRB108 and TRB115/TRB116) and cloned into pBR322 (NEB). From the resulting plasmid DNA, the region was successfully sequenced on both forward and reverse strands.

Potential ORFs were identified using gene prediction tools such as ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/), GeneMark.hmm (47), and Glimmer (48), along with BLAST (31) homology searches and manual annotation. RBSfinder (49) was used to predict ribosome-binding sites (see Table S1 in the supplemental material). ΦM1 tRNAs were identified using tRNAScan-SE (50). The BDGP Neural Network Promoter Prediction (51) program did not identify any consensus promoters. The program Stretcher, from the EMBOSS suite (http://www.ebi.ac.uk/Tools/psa/emboss_stretcher/nucleotide.html), was used for global nucleotide alignments. The ΦM1 genome was viewed and annotated using Artemis (52).

**Plasmid construction.** Molecular biology techniques were performed as described previously (53). All primers were obtained from Sigma-Genosys and Invitrogen and are listed in Table 4. All plasmids constructed and/or used in this study are listed in Table 5, along with the primers used for their construction. All recombinant plasmid sequences were verified by DNA sequencing.

**Measuring $ToxI_{Pa}$ and $ToxN_{Pa}$ levels during phage infection.** Two cultures of 180 ml of LB containing Ap were inoculated with 2-ml overnight cultures of *P. atrosepticum*(pBR322) or *P. atrosepticum*(pMJ4), respectively. Cultures were grown at 25°C and shaken at 180 rpm to an optical density at 600 nm ($OD_{600}$) of 1, and each was split into two 80-ml volumes, one of which was infected with phage at a multiplicity of infection (MOI) of 1, while the other served as a negative control without infection. Cultures were left for 10 min without shaking for phage adsorption and then shaken at 25°C and 180 rpm. Samples for $OD_{600}$ measurement, RNA preparation, and protein analysis were taken regularly during infection. Total RNA was isolated using the TRIzol method and subsequently DNase treated. Cell pellets for Western blot analysis were resuspended in 1× phosphate-buffered saline (PBS) according to the $OD_{600}$ measurement.

**Western blot analysis of $ToxN_{Pa}$ during infection.** One-milliliter samples of the cell cultures were taken, pelleted, and resuspended in 1× PBS according to the $OD_{600}$. For samples taken during ΦM1 phage infection, the protein was quantified using a NanoDrop spectrophotometer (ThermoScientific), and equal amounts of protein (150 $\mu$g) were resolved by 12% PAGE. Proteins were transferred to a polyvinylidene difluoride (PVDF) membrane and blocked for 1 h in 1× PBS containing 5% milk powder. Immunodetection of FLAG-tagged ToxN was performed overnight at 4°C in 1× PBS using anti-FLAG M2 antibody (Sigma). Goat anti-mouse IgG-horseradish peroxidase (HRP) (Santa Cruz) was used as a secondary antibody. Bands were visualized on X-ray film using the SuperSignal West Pico chemilumi-

**TABLE 5** Plasmids used in this study

| Name | Description | Construction source or primers | Template | Resistance |
|---|---|---|---|---|
| pACYC184 | Cloning vector | 59 | | Cm |
| pBR322 | *E. coli* cloning vector | NEB | | Ap, Tc |
| pFR2 | *Photorhabdus luminescens* TT01 full TenpIN$_{Pl}$ locus | 23 | pBR322 | Ap |
| pKNG-uvrA | UvrA marker exchange construct | UvrA upF, UvrA upR, UvrA dnF, UvrA dnR | pKNG101 | Tc, Kan |
| pKNG101-Tc$^r$ | Marker exchange suicide vector | 60 | | Tc |
| pMAT7 | SdhE-FLAG expression vector | 54 | pBAD30 | Ap |
| pMJ4 | ToxI$_{Pa}$, ToxN$_{Pa}$-FLAG with native promoter in pBR322 | 29 | pBR322 | Ap |
| pQE80L | Protein expression vector | Qiagen | | Ap |
| pRW50 | Promoterless LacZ | 38 | | Tc |
| pTA46 | ToxIN$_{Pa}$ with native promoter | 29 | pBR322 | Ap |
| pTA104 | ToxIN$_{Pa}$ promoter | 22 | pRW50 | Tc |
| pTA110 | *In vitro* transcription vector for antisense ToxI$_{Pa}$ RNA | PF217, PF218 | pBSII SK$^-$ | Ap |
| pTRB18-KP14 | ToxI$_{Pa}$ containing | KDO1, PF147 | pACYC184 | Cm, Tc |
| pTRB14 | ToxN$_{Pa}$ CBD | TRB37, TRB38 | pTA46 | Ap |
| pTRB113 | ΦM1 wt construct 3 | TRB126, TRB118 | pBAD30 | Ap, glu |
| pTRB114 | ΦM1 wt construct 4 | TRB117, TRB127 | pBAD30 | Ap, glu |
| pTRB115 | ΦM1 wt construct 5 | TRB126, TRB125 | pBAD30 | Ap, glu |
| pTRB116 | ΦM1 wt construct 6 | TRB128, TRB118 | pBAD30 | Ap, glu |
| pTRB121 | ΦM1-B construct 2 | TRB117, TRB125 | pBAD30 | Ap, glu |
| pTRB123 | ΦM1-B construct 4 | TRB117, TRB127 | pBAD30 | Ap, glu |
| pTRB124 | ΦM1-B construct 5 | TRB126, TRB125 | pBAD30 | Ap, glù |
| pTRB133 | ΦM1 wt construct 7 | TRB111, TRB125 | pBAD30 | Ap, glu |
| pTRB134 | ΦM1 wt construct 8 | TRB129, TRB125 | pBAD30 | Ap, glu |
| pTRB135 | ΦM1 wt construct 9 | TRB130, TRB125 | pBAD30 | Ap, glu |
| pTRB136 | ΦM1-A construct 7 | TRB111, TRB125 | pBAD30 | Ap, glu |
| pTRB139 | ΦM1-B construct 7 | TRB111, TRB125 | pBAD30 | Ap, glu |
| pTRB140 | ΦM1-B construct 8 | TRB129, TRB125 | pBAD30 | Ap, glu |
| pTRB141 | ΦM1-B construct 9 | TRB130, TRB125 | pBAD30 | Ap, glu |
| pTRB148 | ΦM1 wt construct 7-FLAG | TRB111, TRB134 | pBAD30 | Ap, glu |
| pTRB151 | ΦM1-O construct 7-FLAG | TRB111, TRB134 | pBAD30 | Ap, glu |
| pTRB153 | ΦM1-W construct 7-FLAG | TRB111, TRB134 | pBAD30 | Ap, glu |
| pTRB154 | ΦM1-Y construct 7-FLAG | TRB111, TRB134 | pBAD30 | Ap, glu |
| pTRB155 | ΦM1-D construct 7 | TRB111, TRB125 | pBAD30 | Ap, glu |
| pTRB156 | ΦM1-O construct 7 | TRB111, TRB125 | pBAD30 | Ap, glu |
| pTRB157 | ΦM1-V construct 7 | TRB111, TRB125 | pBAD30 | Ap, glu |
| pTRB158 | ΦM1-W construct 7 | TRB111, TRB125 | pBAD30 | Ap, glu |
| pTRB159 | ΦM1-Y construct 7 | TRB111, TRB125 | pBAD30 | Ap, glu |
| pTRB160 | ΦM1 wt LacZ fusion construct | TRB117, TRB127 | pRW50 | Tc |
| pTRB161 | ΦM1 wt LacZ fusion construct | TRB111, TRB127 | pRW50 | Tc |
| pTRB162 | ΦM1 wt LacZ fusion construct | TRB126, TRB127 | pRW50 | Tc |
| pTRB163 | ΦM1-O LacZ fusion construct | TRB117, TRB125 | pRW50 | Tc |
| pTRB164 | ΦM1 wt LacZ fusion construct | TRB117, TRB125 | pRW50 | Tc |
| pTRB189 | ΦM1-23-6His | TRB111, TRB135 | pQE-80L | Ap |
| pTRB190 | ΦM1-O-23-6His | TRB111, TRB135 | pQE-80L | Ap |
| pTRB300 | UvrA-FLAG | TRB330, TRB332 | pBAD33 | Cm, glu |
| pTRB301 | UvrA-6His | TRB337, TRB338 | pQE-80L | Ap |

nescent substrate kit (Pierce). SdhE-FLAG expressed from pMAT7 (54) was used as a control in the blot tracking ΦM1 infection.

**S1 nuclease protection assays.** An antisense probe covering the complete ToxI$_{Pa}$ sequence was made by amplification of the ToxI$_{Pa}$ locus from plasmid pTA110, using primers PF217 and PF218, and subsequent *in vitro* transcription and gel extraction of the probe as described previously (55), generating a uniformly [$^{32}$P]UTP-labeled antisense transcript. Ten micrograms of DNase-treated total RNA was hybridized to the antisense probe overnight at 68°C in a total volume of 30 $\mu$l containing 22% or 6% formamide for the ΦM1 or ΦM1-O total RNA, respectively, 40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)]-KOH (pH 6.4), 1 mM EDTA, and 400 mM NaCl. Reaction mixtures were treated with S1 nuclease (Invitrogen) (1 U $\mu$l$^{-1}$) for 1.5 h at 37°C in a total volume of 300 $\mu$l 1× S1 nuclease buffer to degrade any single-stranded nucleic acids. Double-stranded hybridization products were precipitated, resuspended, and resolved by 10% PAGE. Bands were visualized by phosphorimaging (Bio-Rad Personal FX phosphorimager).

**Toxicity assays.** When required, media were supplemented with Ap, D-glucose (glu) at 0.2% (wt/vol), and L-arabinose (L-ara) at 0.1% (wt/vol). *P. atrosepticum* strains containing two plasmids were grown as 10-ml overnight cultures, used to inoculate 25 ml LB, Ap, Cm, and glu in 250-ml conical flasks, and grown at 25°C and 250 rpm, from a starting OD$_{600}$ of ~0.04, until exponential phase (~1 × 10$^8$ CFU ml$^{-1}$).

Samples were removed, washed with PBS, serially diluted, and plated for viable counts at 25°C on LBA-Ap-Cm plates containing either (i) glu, to repress expression, or (ii) L-ara, to induce expression. Single-plasmid strains were treated in the same way, except that Cm was omitted from the growth conditions.

**β-Galactosidase assays.** Liquid assays for LacZ activity were performed using the substrate 4′-methylumbelliferyl-β-D-glucuronide (MUG) as described before (56). Briefly, samples of culture (150 μl) were taken at each time point and frozen at −80°C until required. Ten-microliter aliquots of each sample culture were frozen at −80°C for 10 min and then thawed at room temperature. Next, 100 μl reaction buffer (PBS, 400 μg $ml^{-1}$ lysozyme, 250 μg $ml^{-1}$ MUG) was added, and samples were immediately monitored in a Gemini XPS plate reader with the following parameters: excitation, 360 nm; emission, 450 nm; cutoff, 435 nm; eight reads per well; measurements taken every 30 s for 30 min. Relative fluorescence units $min^{-1}$ was calculated from a period of linear increase in fluorescence, normalized to the $OD_{600}$ of the sample.

**Pulldown of $ToxIN_{Pa}$ and M1-23 from cell lysates.** Using ΦM1 and ΦM1-O genomic DNA, ΦM1-23 and M1-O-23 were amplified via PCR using TRB111 and TRB135 as primers. The products were then digested using the relevant restriction enzymes, ligated into pQE-80L, and then used to transform ER2566. For the $ToxIN_{Pa}$ strains, pMJ4 (which contains $ToxIN_{Pa}$-FLAG) was used, and a new plasmid was constructed to make a $ToxN_{Pa}$-chitin binding domain (CBD) fusion. This was produced using pTA46 and primers TRB37 and TRB38. The plasmid pTRB14 was then used to transform ER2566, which had previously been transformed with pTRB18-KP14, which contains a $ToxI_{Pa}$ sequence.

Expression strains were grown in 2× YT medium (per liter, 16 g tryptone, 10 g yeast extract, 5 g NaCl) at 37°C until an $OD_{600}$ of approximately 1. The cultures were then induced with the appropriate supplement (0.5 mM IPTG for M1-23-6His and M1-O-23-6His) and then left to grow overnight at 18°C. No inducers were added to the tagged $ToxIN_{Pa}$-containing strains, as $ToxIN_{Pa}$ is constitutively expressed on pBR322.

Cells were harvested by centrifugation at 8,000 × *g*, and the pellets were resuspended in 10 ml lysis buffer (50 mM $NaH_2PO_4{\cdot}2H_2O$, 500 mM NaCl, 10 mM imidazole, 10% glycerol, pH 8.0) per 500 ml of original culture volume. Cells were then lysed by four passes through a high-pressure homogenizer (EmulsiFlex; at up to 15,000 $lb/in^2$). Lysed cells were centrifuged at 8,000 × *g*, and the supernatants were kept for further coimmunoprecipitation experiments.

In the experiments using M1-23-6His and M1-O-23-6His as bait, 1.5 ml $Ni^{2+}$ resin columns were used with $ToxIN_{Pa}$-FLAG. The columns were equilibrated using 3 column volumes (CV) of lysis buffer before the His-tagged protein lysates were loaded onto the resin. Loaded resins were washed with 5 CV of wash buffer 1 (20 mM imidazole), followed by 10 CV of wash buffer 2 (40 mM imidazole). The FLAG-tagged $ToxIN_{Pa}$ was then loaded onto the appropriate columns via continuous flow for at least 3 h (often overnight) before washing with 5 CV wash buffer 1 and 10 CV wash buffer 2.

Samples were eluted from the resin using elution buffer (250 mM imidazole) via three 1-ml fractions and analyzed by Western blot analysis using antibodies against His (Novagen) and FLAG (Sigma) tags. Briefly, samples were run on 12.5% Tris-Tricine gels and transferred onto Immobilon-P PVDF membranes (pore size, 0.45 μm; Millipore) at 250 mA for 90 min. Membranes were then blocked with a 5% milk plus PBST (PBS with Tween 20) solution for 1 h before incubation with anti-His and anti-FLAG antibodies at 1:10,000 for 2 h. After incubation, the membranes were washed three times for 5 min each in PBST and then incubated with the secondary anti-mouse antibody (Sigma) at 1:10,000 for 1 h before they were washed again three times for 5 min each in PBST. The blots were then probed with Immobilon-Western chemiluminescent HRP-substrate (Millipore) and developed.

For experiments in which $ToxIN_{Pa}$ was used as the bait, the strain expressing $ToxIN_{Pa}$-CBD was used with 1 ml chitin resin. The protocol and buffers used were as described by the manufacturer (NEB). Briefly, the $ToxIN_{Pa}$-CBD lysate was loaded onto the column and washed with 40 ml of column buffer. The M1-23 or control pQE-80L lysates were then added to their respective columns. The columns were washed twice with 10 ml and then 27 ml of column buffer, followed by a dithiothreitol (DTT) flush, 5 to 7 ml for 10 min. Columns were then left to incubate overnight at room temperature. After incubation, elution was carried out using 15 ml of column buffer. Western blot analyses were then performed on the samples as previously described.

**Measuring $ToxI_{Pa}$ levels after $ToxIN_{Pa}$ pulldown with M1-23.** $ToxI_{Pa}$ levels were measured in the eluted fractions of the $ToxIN_{Pa}$-CBD chitin resin column experiments. Samples from cultures either expressing M1-23 or containing the pQE-80L vector control were separated by electrophoresis at 80 V, using a 1% (wt/vol) agarose gel made with 0.5× TAE (Tris-acetate-EDTA). Additionally, samples were also measured with a NanoDrop spectrophotometer (Labtech; ND-1000).

**Coimmunoprecipitation of UvrA and M1-23.** UvrA-6His was constructed by amplification from the *E. coli* W3110 genome using primers TRB337 and TRB338. PCR products were then digested with the appropriate restriction enzymes, and the digested product was purified and then ligated into pQE-80L to generate UvrA with an N-terminal His tag, pTRB301. This plasmid was then used to transform the *E. coli* expression strain ER2566. Likewise, UvrA-FLAG was constructed in a similar way but using primers TRB330 and TRB332 and ligated into pBAD33.

Expression and subsequent experiments were performed as described earlier using His-tagged proteins as bait on $Ni^{2+}$ resin. Expression of UvrA-FLAG was induced by the addition of 0.02% arabinose.

**Construction of the *P. atrosepticum uvrA* mutant.** The *uvrA* mutant of *P. atrosepticum* was constructed via allelic exchange. This was performed using the plasmid pKNG-uvrA, which was derived from pKNG101. The plasmid was constructed by first amplifying 500-bp regions up- and downstream of

the *uvrA* gene in *P. atrosepticum* SCRI1043. These two sequences were then ligated together with a kanamycin cassette inserted in between.

The suicide vector derivative pKNG-uvrA was used to transform *E. coli* β2163 and grown overnight in the appropriate selective medium. This served as the donor strain and, along with an overnight culture of the recipient strain, *P. atrosepticum* SCRI1043, was pelleted and resuspended in LB. Both cultures were then mixed at the ratios of 2:1, 1:1, and 1:2 up to a final volume of 100 μl. The resulting mixtures were then spotted on DAPA-containing plates and incubated at 25°C for 24 h. After mating, the patches were resuspended in 100 μl LB, serially diluted, and spread onto LBA plates containing tetracycline. These plates were incubated for 2 days at 25°C, and colonies that appeared were picked and grown in LB overnight. The subsequent overnight cultures were serially diluted, and 50-μl samples were plated onto LBA plates containing 10% (wt/vol) sucrose. Colonies were also patched onto LBA plates containing kanamycin, and the gene deletion was confirmed using colony PCR and DNA sequencing. The strain was confirmed phenotypically as UvrA negative by demonstrating a hypersensitivity to UV light (see Fig. S1 in the supplemental material).

**Accession number(s).** The genome of ΦM1 has been submitted to GenBank under the accession number JX290549.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.03229-16.

**SUPPLEMENTAL FILE 1,** PDF file, 1.0 MB.

## ACKNOWLEDGMENTS

This research was funded by support from the Biotechnology and Biological Sciences Research Council, United Kingdom, to G.P.C.S., by a Marsden Fund, Royal Society of New Zealand (RSNZ) award to P.C.F. and G.P.C.S., and by a Rutherford Discovery Fellowship (RSNZ) to P.C.F. B.F.L. was supported by the Wellcome Trust, United Kingdom. Work carried out with *P. atrosepticum* was under DEFRA license no. 50864/197900/3.

We also acknowledge Simon Poulter for the construction of pKNG101-Tc$^r$ (Table 5).

## REFERENCES

1. Wommack KE, Colwell RR. 2000. Virioplankton: viruses in aquatic ecosystems. Microbiol Mol Biol Rev 64:69–114. https://doi.org/10.1128/MMBR.64.1.69-114.2000.
2. Chibani-Chennoufi S, Bruttin A, Dillmann ML, Brüssow H. 2004. Phage-host interaction: an ecological perspective. J Bacteriol 186:3677–3686. https://doi.org/10.1128/JB.186.12.3677-3686.2004.
3. Dy RL, Richter C, Salmond GP, Fineran PC. 2014. Remarkable mechanisms in microbes to resist phage infections. Annu Rev Virol 1:307–331. https://doi.org/10.1146/annurev-virology-031413-085500.
4. Stern A, Sorek R. 2011. The phage-host arms race: shaping the evolution of microbes. Bioessays 33:43–51. https://doi.org/10.1002/bies.201000071.
5. Loenen WA, Dryden DT, Raleigh EA, Wilson GG, Murray NE. 2014. Highlights of the DNA cutters: a short history of the restriction enzymes. Nucleic Acids Res 42:3–19. https://doi.org/10.1093/nar/gkt990.
6. Doudna JA, Charpentier E. 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346:1258096. https://doi.org/10.1126/science.1258096.
7. Chopin MC, Chopin A, Bidnenko E. 2005. Phage abortive infection in lactococci: variations on a theme. Curr Opin Microbiol 8:473–479. https://doi.org/10.1016/j.mib.2005.06.006.
8. Cavanagh D, Fitzgerald GF, McAuliffe O. 2015. From field to fermentation: the origins of *Lactococcus lactis* and its domestication to the dairy environment. Food Microbiol 47:45–61. https://doi.org/10.1016/j.fm.2014.11.001.
9. Parma DH, Snyder M, Sobolevski S, Nawroz M, Brody E, Gold L. 1992. The Rex system of bacteriophage lambda: tolerance and altruistic cell death. Genes Dev 6:497–510. https://doi.org/10.1101/gad.6.3.497.
10. Georgiou T, Yu YN, Ekunwe S, Buttner MJ, Zuurmond A, Kraal B, Kleanthous C, Snyder L. 1998. Specific peptide-activated proteolytic cleavage of Escherichia coli elongation factor Tu. Proc Natl Acad Sci U S A 95:2891–2895. https://doi.org/10.1073/pnas.95.6.2891.
11. Penner M, Morad I, Snyder L, Kaufmann G. 1995. Phage T4-coded Stp: double-edged effector of coupled DNA and tRNA-restriction systems. J Mol Biol 249:857–868. https://doi.org/10.1006/jmbi.1995.0343.
12. Yamaguchi Y, Park JH, Inouye M. 2011. Toxin-antitoxin systems in bacteria and archaea. Annu Rev Genet 45:61–79. https://doi.org/10.1146/annurev-genet-110410-132412.
13. Gerdes K, Rasmussen PB, Molin S. 1986. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. Proc Natl Acad Sci U S A 83:3116–3120. https://doi.org/10.1073/pnas.83.10.3116.
14. Ramage HR, Connolly LE, Cox JS. 2009. Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. PLoS Genet 5:e1000767. https://doi.org/10.1371/journal.pgen.1000767.
15. Christensen SK, Gerdes K. 2003. RelE toxins from Bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. Mol Microbiol 48:1389–1400. https://doi.org/10.1046/j.1365-2958.2003.03512.x.
16. Lehnherr H, Maguin E, Jafri S, Yarmolinsky MB. 1993. Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. J Mol Biol 233:414–428. https://doi.org/10.1006/jmbi.1993.1521.
17. Page R, Peti W. 2016. Toxin-antitoxin systems in bacterial growth arrest and persistence. Nat Chem Biol 12:208–214. https://doi.org/10.1038/nchembio.2044.
18. Brantl S, Jahn N. 2015. sRNAs in bacterial type I and type III toxin-antitoxin systems. FEMS Microbiol Rev 39:413–427. https://doi.org/10.1093/femsre/fuv003.
19. Pecota DC, Wood TK. 1996. Exclusion of T4 phage by the hok/sok killer locus from plasmid R1. J Bacteriol 178:2044–2050. https://doi.org/10.1128/jb.178.7.2044-2050.1996.
20. Hazan R, Engelberg-Kulka H. 2004. *Escherichia coli* mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1. Mol Genet Genomics 272:227–234.

21. Otsuka Y, Yonesaki T. 2012. Dmd of bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnlA toxins. Mol Microbiol 83:669–681. https://doi.org/10.1111/j.1365-2958.2012.07975.x.
22. Fineran PC, Blower TR, Foulds IJ, Humphreys DP, Lilley KS, Salmond GP. 2009. The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. Proc Natl Acad Sci U S A 106:894–899. https://doi.org/10.1073/pnas.0808832106.
23. Blower TR, Short FL, Rao F, Mizuguchi K, Pei XY, Fineran PC, Luisi BF, Salmond GP. 2012. Identification and classification of bacterial type III toxin-antitoxin systems encoded in chromosomal and plasmid genomes. Nucleic Acids Res 40:6158–6173. https://doi.org/10.1093/nar/gks231.
24. Samson JE, Spinelli S, Cambillau C, Moineau S. 2013. Structure and activity of AbiQ, a lactococcal endoribonuclease belonging to the type III toxin-antitoxin system. Mol Microbiol 87:756–768. https://doi.org/10.1111/mmi.12129.
25. Dy RL, Przybilski R, Semeijn K, Salmond GP, Fineran PC. 2014. A widespread bacteriophage abortive infection system functions through a type IV toxin-antitoxin mechanism. Nucleic Acids Res 42:4590–4605. https://doi.org/10.1093/nar/gkt1419.
26. Sberro H, Leavitt A, Kiro R, Koh E, Peleg Y, Qimron U, Sorek R. 2013. Discovery of functional toxin/antitoxin systems in bacteria by shotgun cloning. Mol Cell 50:136–148. https://doi.org/10.1016/j.molcel.2013.02.002.
27. Blower TR, Evans TJ, Przybilski R, Fineran PC, Salmond GP. 2012. Viral evasion of a bacterial suicide system by RNA-based molecular mimicry enables infectious altruism. PLoS Genet 8:e1003023. https://doi.org/10.1371/journal.pgen.1003023.
28. Samson JE, Bélanger M, Moineau S. 2013. Effect of the abortive infection mechanism and type III toxin/antitoxin system AbiQ on the lytic cycle of *Lactococcus lactis* phages. J Bacteriol 195:3947–3956. https://doi.org/10.1128/JB.00296-13.
29. Blower TR, Fineran PC, Johnson MJ, Toth IK, Humphreys DP, Salmond GP. 2009. Mutagenesis and functional characterization of the RNA and protein components of the toxIN abortive infection and toxin-antitoxin locus of *Erwinia*. J Bacteriol 191:6029–6039. https://doi.org/10.1128/JB.00720-09.
30. Toth IK, Mulholland V, Cooper V, Bentley S, Shih Y, Perombelon MCM, Salmond GPC. 1997. Generalized transduction in the potato blackleg pathogen *Erwinia carotovora* subsp. *atroseptica* by bacteriophage phi M1. Microbiology 143:2433–2438. https://doi.org/10.1099/00221287-143-7-2433.
31. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
32. Ceyssens PJ, Lavigne R, Mattheus W, Chibeu A, Hertveldt K, Mast J, Robben J, Volckaert G. 2006. Genomic analysis of *Pseudomonas aeruginosa* phages LKD16 and LKA1: establishment of the phiKMV subgroup within the T7 supergroup. J Bacteriol 188:6924–6931. https://doi.org/10.1128/JB.00831-06.
33. Lavigne R, Burkal'tseva MV, Robben J, Sykilinda NN, Kurochkina LP, Grymonprez B, Jonckx B, Krylov VN, Mesyanzhinov VV, Volckaert G. 2003. The genome of bacteriophage phiKMV, a T7-like virus infecting *Pseudomonas aeruginosa*. Virology 312:49–59. https://doi.org/10.1016/S0042-6822(03)00123-5.
34. Adriaenssens EM, Van Vaerenbergh J, Vandenheuvel D, Dunon V, Ceyssens PJ, De Proft M, Kropinski AM, Noben JP, Maes M, Lavigne R. 2012. T4-related bacteriophage LIMEstone isolates for the control of soft rot on potato caused by '*Dickeya solani*'. PLoS One 7:e33227. https://doi.org/10.1371/journal.pone.0033227.
35. Bell KS, Sebaihia M, Pritchard L, Holden MT, Hyman LJ, Holeva MC, Thomson NR, Bentley SD, Churcher LJ, Mungall K, Atkin R, Bason N, Brooks K, Chillingworth T, Clark K, Doggett J, Fraser A, Hance Z, Hauser H, Jagels K, Moule S, Norbertczak H, Ormond D, Price C, Quail MA, Sanders M, Walker D, Whitehead S, Salmond GP, Birch PR, Parkhill J, Toth IK. 2004. Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. Proc Natl Acad Sci U S A 101:11105–11110. https://doi.org/10.1073/pnas.0402424101.
36. Bastías R, Higuera G, Sierralta W, Espejo RT. 2010. A new group of cosmopolitan bacteriophages induce a carrier state in the pandemic strain of *Vibrio parahaemolyticus*. Environ Microbiol 12:990–1000. https://doi.org/10.1111/j.1462-2920.2010.02143.x.
37. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177:4121–4130. https://doi.org/10.1128/jb.177.14.4121-4130.1995.
38. Lodge J, Fear J, Busby S, Gunasekaran P, Kamini NR. 1992. Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. FEMS Microbiol Lett 74:271–276.
39. Short FL, Pei XY, Blower TR, Ong SL, Fineran PC, Luisi BF, Salmond GP. 2013. Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot. Proc Natl Acad Sci U S A 110:E241–E249. https://doi.org/10.1073/pnas.1216039110.
40. Little JW, Mount DW. 1982. The SOS regulatory system of *Escherichia coli*. Cell 29:11–22. https://doi.org/10.1016/0092-8674(82)90085-X.
41. Wagner EG, Unoson C. 2012. The toxin-antitoxin system tisB-istR1: expression, regulation, and biological role in persister phenotypes. RNA Biol 9:1513–1519. https://doi.org/10.4161/rna.22578.
42. Maisonneuve E, Castro-Camargo M, Gerdes K. 2013. (p)ppGpp controls bacterial persistence by stochastic induction of toxin-antitoxin activity. Cell 154:1140–1150. https://doi.org/10.1016/j.cell.2013.07.048.
43. Verstraeten N, Knapen WJ, Kint CI, Liebens V, Van den Bergh B, Dewachter L, Michiels JE, Fu Q, David CC, Fierro AC, Marchal K, Beirlant J, Versées W, Hofkens J, Jansen M, Fauvart M, Michiels J. 2015. Obg and membrane depolarization are part of a microbial bet-hedging strategy that leads to antibiotic tolerance. Mol Cell 59:9–21. https://doi.org/10.1016/j.molcel.2015.05.011.
44. Kalischuk M, Hachey J, Kawchuk L. 2015. Complete genome sequence of phytopathogenic *Pectobacterium atrosepticum* bacteriophage Peat1. Genome Announc 3:e00760–15. https://doi.org/10.1128/genomeA.00760-15.
45. Hirata H, Kashihara M, Horiike T, Suzuki T, Dohra H, Netsu O, Tsuyumu S. 2016. Genome sequence of *Pectobacterium carotovorum* phage PPWS1, isolated from Japanese horseradish [Eutrema japonicum (Miq.) Koidz] showing soft-rot symptoms. Genome Announc 4:e01625–15. https://doi.org/10.1128/genomeA.01625-15.
46. Petty NK, Foulds IJ, Pradel E, Ewbank JJ, Salmond GP. 2006. A generalized transducing phage (phiIF3) for the genomically sequenced *Serratia marcescens* strain Db11: a tool for functional genomics of an opportunistic human pathogen. Microbiology 152:1701–1708. https://doi.org/10.1099/mic.0.28712-0.
47. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res 26:1107–1115. https://doi.org/10.1093/nar/26.4.1107.
48. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27:4636–4641. https://doi.org/10.1093/nar/27.23.4636.
49. Suzek BE, Ermolaeva MD, Schreiber M, Salzberg SL. 2001. A probabilistic method for identifying start codons in bacterial genomes. Bioinformatics 17:1123–1130. https://doi.org/10.1093/bioinformatics/17.12.1123.
50. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964. https://doi.org/10.1093/nar/25.5.0955.
51. Reese MG. 2001. Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. Comput Chem 26:51–56. https://doi.org/10.1016/S0097-8485(01)00099-7.
52. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. Bioinformatics 16:944–945. https://doi.org/10.1093/bioinformatics/16.10.944.
53. Fineran PC, Everson L, Slater H, Salmond GP. 2005. A GntR family transcriptional regulator (PigT) controls gluconate-mediated repression and defines a new, independent pathway for regulation of the tripyrrole antibiotic, prodigiosin, in *Serratia*. Microbiology 151:3833–3845. https://doi.org/10.1099/mic.0.28251-0.
54. McNeil MB, Clulow JS, Wilf NM, Salmond GP, Fineran PC. 2012. SdhE is a conserved protein required for flavinylation of succinate dehydrogenase in bacteria. J Biol Chem 287:18418–18428. https://doi.org/10.1074/jbc.M111.293803.
55. Przybilski R, Richter C, Gristwood T, Clulow JS, Vercoe RB, Fineran PC. 2011. Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. RNA Biol 8:517–528. https://doi.org/10.4161/rna.8.3.15190.
56. Ramsay JP, Williamson NR, Spring DR, Salmond GP. 2011. A quorum-sensing molecule acts as a morphogen controlling gas vesicle organelle biogenesis and adaptive flotation in an enterobacterium. Proc

Natl Acad Sci U S A 108:14932–14937. https://doi.org/10.1073/pnas.1109169108.
57. Demarre G, Guérout AM, Matsumoto-Mashimo C, Rowe-Magnus DA, Marlière P, Mazel D. 2005. A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPα) conjugative machineries and their cognate *Escherichia coli* host strains. Res Microbiol 156:245–255. https://doi.org/10.1016/j.resmic.2004.09.007.
58. Bachmann BJ. 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol Rev 36:525–557.
59. Chang AC, Cohen SN. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J Bacteriol 134:1141–1156.
60. Poulter S. 2011. The LuxR-family quorum sensing transcriptional regulator CarR in *Erwinia* and *Serratia*. PhD thesis. University of Cambridge, Cambridge, United Kingdom.